

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Sciences cognitives, Psychologie et Neurocognition**

Arrêté ministériel : 7 août 2006

Présentée par

Krystyna GRABSKI

Thèse dirigée par **Marc SATO** et **Jean-Luc SCHWARTZ**

préparée au sein du **Département Parole & Cognition du Gipsa-lab,**
UMR CNRS 5216 & Université de Grenoble,

dans l'**École Doctorale Ingénierie pour la Santé, la Cognition et**
l'Environnement

Les cartes sensorimotrices de la parole : Corrélat neurocognitifs et couplage fonctionnel des systèmes de perception et de production des voyelles du Français.

Thèse soutenue publiquement le **27 septembre 2012,**

devant le jury composé de :

Madame Monica Baciú, Professeur, Université Pierre Mendès-France,
Laboratoire de Psychologie et NeuroCognition, UMR CNRS 5105, France -
Examineur

Monsieur Jean-François Démonet, Directeur de Recherche INSERM,
Professeur, Université de Lausanne, Centre Leenaards de la Mémoire, Centre
Hospitalier Universitaire Vaudois, Suisse - Rapporteur

Monsieur Luciano Fadiga, Professeur, Università degli studi di Ferrara,
Dipartimento di Scienze Biomediche e Terapie Avanza, Istituto Italiano di
Tecnologia, Italie - Examineur

Monsieur Noël Nguyen, Professeur, Aix-Marseille Université, Laboratoire Parole
& Langage, UMR CNRS 6057, France - Rapporteur

Monsieur Jean-Luc Schwartz, Directeur de Recherche CNRS, Université de
Grenoble, Gipsa-lab UMR CNRS 5216, France - Co-Directeur de thèse

Monsieur Marc Sato, Chargé de Recherche CNRS, Université de Grenoble, Gipsa-
lab UMR CNRS 5216, France - Directeur de thèse



THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Sciences cognitives, Psychologie et Neurocognition**

Arrêté ministériel : 7 août 2006

Présentée par

Krystyna GRABSKI

Thèse dirigée par **Marc SATO** et **Jean-Luc SCHWARTZ**

préparée au sein du **Département Parole & Cognition du Gipsa-lab,**
UMR CNRS 5216 & Université de Grenoble,

dans l'**École Doctorale Ingénierie pour la Santé, la Cognition et**
l'Environnement

Les cartes sensorimotrices de la parole : Corrélat neurocognitifs et couplage fonctionnel des systèmes de perception et de production des voyelles du Français.

Thèse soutenue publiquement le **27 septembre 2012,**

devant le jury composé de :

Madame Monica Baciú, Professeur, Université Pierre Mendès-France,
Laboratoire de Psychologie et NeuroCognition, UMR CNRS 5105, France -
Examineur

Monsieur Jean-François Démonet, Directeur de Recherche INSERM,
Professeur, Université de Lausanne, Centre Leenaards de la Mémoire, Centre
Hospitalier Universitaire Vaudois, Suisse - Rapporteur

Monsieur Luciano Fadiga, Professeur, Università degli studi di Ferrara,
Dipartimento di Scienze Biomediche e Terapie Avanza, Istituto Italiano di
Tecnologia, Italie - Examineur

Monsieur Noël Nguyen, Professeur, Aix-Marseille Université, Laboratoire Parole
& Langage, UMR CNRS 6057, France - Rapporteur

Monsieur Jean-Luc Schwartz, Directeur de Recherche CNRS, Université de
Grenoble, Gipsa-lab UMR CNRS 5216, France - Co-Directeur de thèse

Monsieur Marc Sato, Chargé de Recherche CNRS, Université de Grenoble, Gipsa-
lab UMR CNRS 5216, France - Directeur de thèse



Résumé/Abstract

LES CARTES SENSORIMOTRICES DE LA PAROLE : Corrélats neurocognitifs et couplage fonctionnel des systèmes de perception et de production des voyelles du Français

La parole est construite sur un jeu de correspondances entre représentations sensorielles et articulatoires, notamment lors de l'acquisition du langage les premières années de vie. Par l'utilisation de l'imagerie par résonance magnétique fonctionnelle, l'objectif premier de nos travaux était de déterminer, chez l'adulte, un possible couplage fonctionnel des systèmes de perception et de production des voyelles du Français, considérées comme unités élémentaires de la parole. En parallèle, nos travaux devaient permettre de clarifier les structures cérébrales liées au contrôle moteur orofacial de mouvements simples supralaryngés et, à l'aide de la technique de stimulation magnétique transcrânienne, de déterminer une possible implication causale des régions sensorielles et motrices lors de la perception de la parole. Nos travaux ont permis de souligner l'implication des régions sensorielles et motrices aussi bien lors de la réalisation des gestes orofaciaux que lors de la production et de la perception des voyelles. La mise en évidence d'un effet d'adaptation pour ces régions motrices, auditives et somatosensorielles lors de l'écoute ou de la réalisation répétée d'une même voyelle ou d'un même geste suggère de plus l'existence de boucles sensorimotrices communes, impliquées dans des mécanismes adaptatifs de contrôle en ligne des gestes de parole perçus et produits. Enfin, nous avons pu démontrer le rôle causal et fonctionnel des régions sensorielles et motrices de la voie dorsale lors de la catégorisation de sons de parole. Pris ensemble, nos travaux soulignent la nature distribuée sensorimotrice des représentations cérébrales des unités de parole.

Mots clés : perception et production de la parole, voyelles, contrôle moteur orofacial, interactions sensorimotrices, représentations et cartes neurocognitives, IRMf, TMS.

SENSORIMOTOR MAPS OF SPEECH: Neurocognitive correlates and functional coupling of French vowel perception and production systems

Speech is built on a set of correspondences between sensory and articulatory representations, especially during the acquisition of language in the early years of life. Using functional magnetic resonance imaging, the primary goal of our work was to determine, in adults, a possible functional coupling of French vowel perception and production systems, as elementary speech units. In parallel, our work should help to clarify the brain structures related to the orofacial motor control of simple supralaryngeal movements and to determine a possible causal contribution of sensory and motor regions during speech perception. Our work highlights the involvement of sensory and motor areas when performing orofacial gestures and during vowel production and perception. Adaptive effects of these motor, auditory and somatosensory regions during repeated orofacial movements and in vowel perception and production suggest the existence of common adaptive mechanisms involved in the online control of perceived and produced speech gestures. Finally, we demonstrated a causal and functional role of the sensorimotor regions of the dorsal pathway in speech categorization. Taken together, our results emphasize the distributed sensorimotor nature of cerebral representations of speech units.

Key words: speech perception and production, vowels, orofacial motor control, sensorimotor interactions, neurocognitive representations and maps, fMRI, TMS.

Remerciements

En premier lieu, je tiens à remercier Marc Sato et Jean-Luc Schwartz de m'avoir acceptée en doctorat et proposé ce sujet passionnant, reliant les univers de la parole et des neurosciences. Je vous suis reconnaissante de m'avoir permis de travailler dans le Département Parole & Cognition, lieu extraordinaire pour découvrir le monde multi- et interdisciplinaire de la parole et de m'avoir ouvert la voie de la neurophonétique. Vous m'avez tellement appris et j'ai encore tellement à apprendre de vous.

Marc, *wielkie podziękowania* tout particulièrement. Il est difficile d'exprimer en quelques mots toute ma gratitude. Merci pour ces années de "belles aventures" scientifiques guidées par ta passion pour la science, ta volonté à transmettre et partager ton enthousiasme, connaissances et savoir-faire - à la fois avec exigence et patience. J'ai conscience qu'étant ta première thésarde, un certain challenge était de mise. Je te remercie d'autant plus pour ta confiance et ta générosité. Ainsi j'ai toujours pu compter sur ton soutien, même et surtout dans les moments difficiles.

Jean-Luc, même si Marc était mon encadrant principal, je n'oublie pas que tu étais co-initiateur de cette thèse. Merci pour toutes les précieuses sagesses sur la parole et aussi pour les paroles de sagesse quand il les fallait, et pour tes encouragements, notamment les derniers mois.

J'adresse mes remerciement à Monica Baciú, Jean-François Démonet, Luciano Fadiga et Noël Nguyen de m'avoir fait l'honneur de participer à mon jury de thèse, pour leur temps consacré à lire, se déplacer et évaluer ce travail. J'aimerais remercier les membres du jury également pour leurs retours encourageants et leurs remarques constructives.

Lors de ma thèse j'ai eu l'opportunité et le très grand plaisir d'enseigner en tant que vacataire et ensuite ATER en neurosciences au Département de Psychologie à l'UPMF. Ici je tiens à remercier particulièrement Monica Baciú pour sa confiance, encouragements et gentillesse, mais également Aurélie Campagne, Adriana Perles, les personnes administratives et de l'accueil tout comme les étudiants, grâce à qui cette expérience me fut précieuse, source d'inspirations et de satisfactions. Ces enseignements m'ont certainement procuré encore plus de plaisir que de travail.

Pendant ma thèse j'ai eu la chance de partir plusieurs fois à Montréal, notamment au CRLMB (devenu CRBLM). Je remercie très chaleureusement Vincent Gracco pour cette rencontre extrêmement enrichissante, pour son accueil, la mise à disposition d'un environnement de travail formidable et également l'opportunité d'apprendre une nouvelle technique d'investigation. Ici j'en profite pour remercier également Pascale Tremblay pour sa collaboration pendant et après mon séjour, mais surtout pour les longues heures passées devant la machine TMS, pour son instruction rigoureuse, son incroyable énergie... tout ça avec tant d'humour. Je suis ravie d'avoir appris autant grâce à vous.

Je remercie toutes les autres personnes qui ont collaboré ou contribué à ces projets, en particulier Nathalie Vallée et Coriandre Vilain du Gipsa-lab, Laurent Lamalle et Jean-François Le Bas de l'IRM 3T, Monica Baciú du LPNC. Merci également à Lionel Granjon, Cédric Pichat et Emilie Cousin pour votre aide précieuse d'ingénieurs; Nadine, Houria et Akila pour l'aide administrative. Sans oublier les participants, grenoblois et montréalais, pour leur courage et disponibilité, ainsi que toutes les personnes m'ayant apporté des échanges ou critiques constructifs.

Une place à part pour les nombreux collègues (permanents et non permanents) au DPC, que je remercie chaleureusement pour leur accueil, amitié et discussions tout au long des quatre années. Grâce à vous, j'ai pu baigner dans un environnement à large spectre qui tourne autour de la parole, découvrir ou creuser des notions entre autres dans les domaines d'interactions sensorimotrices

(merci à Marc Sato), de la phonétique (merci à Nathalie Vallée et Hien Tran), des modèles internes (merci à Pascal Perrier) et tant d'autres questions sur la parole, la cognition et les neurosciences (merci, entre autres, à Louis-Jean Boë, Véronique Aubergé,... mais surtout Jean-Luc Schwartz et Frédéric Berthommier). Ce fut un honneur et un plaisir de travailler parmi et avec vous, tant sur le plan scientifique qu'humain.

Une très chaleureuse pensée à tous les collègues thésards (ou actuellement postdocs, ingénieurs ou permanents) au DPC Stendhal et Ampère pour cette belle complicité (petite pensée à Guillaume pour son optimisme passionné mythique) et particulièrement Anna, Anne, Emilie, Hien, Ibrah, Marylamb, Nico, Paolo, Rosario, Sandra, Silvia... merci pour votre amitié, votre soutien et disponibilité, toutes ces discussions, scientifiques ou de comptoir, les rires.. tout cela me manque déjà. Maëva, ma co-bureau des derniers mois, les plus éprouvants, merci pour ta gentillesse et ta compréhension ! Une bien chaleureuse pensée aussi aux anciens thésards voisins du LPNC ou du GIN, notamment Benoît, Marcela, Muriel... et bien sûr Katell.

Merci aux amis, qui m'ont supporté et ont cru en moi, de près ou de loin. Dans le désordre et de manière non exhaustive : Emilie et Nicolas (pour les folies, les rire, balades, danses nécessaires pour mon équilibre vital), Sophie (pour ta générosité et ton sens de la réalité), Victor (pour ma décade française teintée de Vercors), Quentin (pour la musique). Je pense à Barbara (ma sœur française), Julien, à Joachim, Manu, mes amis-colocataires et les autres de Lyon et Marseille, Yolaine, Elise et particulièrement David et Caro, qui ont cru en moi dès le départ (j'ai souvent pensé à vous en des moments difficiles). Et bien sûr, mes amis polonais, qui bien que loin depuis des années et malgré mes longues absences, sont toujours restés fidèles au poste et présents dans mes pensées. Magda, Kasia i Kasia, Kamila, Piterka, Blo, Jarek, Yarek, Krzysiek, Murdza... et tant d'autres qui se reconnaîtront !

Enfin, je me rends compte de ma chance d'être entourée par ma très belle et grande famille, qui est source d'énergie, de courage et d'inspiration, tant émotionnelle, spirituelle qu'intellectuelle.

A mes parents, Maria et Józef, merci de m'avoir inculqué le goût pour les langues et l'aventure, que rien n'est simple mais pas impossible non plus. Merci d'avoir été là, avec votre foi et soutien infaillible, dès le début.

"Grey matters" - MN/

Misiom, które wierzyły w oczka z guzika

Table des matières

Introduction	1
La perception de la parole sous l'angle des interactions sensorimotrices	4
Approches auditives, théories motrices et sensorimotrices	4
Codage cérébral de la perception de la parole	11
La production de la parole sous l'angle des interactions sensorimotrices	22
Apports des études comportementales	22
Codage cérébral de la production de la parole	26
Problématique de la thèse et guide de lecture	34
 1^{ère} Partie : Les cartes sensorimotrices de la parole	 37
Avant-propos	39
Considérations méthodologiques	40
Protocole expérimental	40
Acquisition des données IRM	42
Chapitre 1.1 : Corrélats neuronaux des gestes orofaciaux	47
Résumé/Abstract	47
1.1.1 Introduction	48
Corrélats neuronaux du contrôle moteur orofacial	48
Cadre de l'étude	50
1.1.2 Méthode	51
Protocole	51
Analyses statistiques	52
1.1.3 Résultats	54
Réseaux neuronaux des mouvements orofaciaux	54
Somatotopie sensorimotrice	56
Activités cérébrales liées aux délais d'acquisition	61
1.1.4 Discussion	62
Réseau orofacial minimal	62
Effet principal de la tâche motrice	63
Somatotopie articulaire	64
Patterns temporels d'activations	65
1.1.5 Conclusion	66
Chapitre 1.2 : Corrélats neuronaux des systèmes de perception et de production des voyelles du Français	67
Résumé/Abstract	67
1.2.1 Introduction	68
Codage des unités de parole	69
Cadre de l'étude	70

1.2.2 Méthode	71
Protocole	71
Analyses statistiques	75
1.2.3 Résultats	76
Perception et production des voyelles	77
Activités cérébrales liées aux délais d'acquisition	83
1.2.4 Discussion	87
Système de production des voyelles	88
Système de perception des voyelles.....	89
Activations différenciées lors de la perception et de la production des voyelles	92
Couplage fonctionnel des systèmes de perception et de production des voyelles.....	93
Représentations différenciées entre voyelles	94
Délais d'acquisition	95
1.2.5 Conclusion.....	96

2^{nde} Partie : Contrôle sensorimoteur adaptatif et codage du but des actions orofaciales et vocaliques

Avant-propos	101
--------------------	-----

Chapitre 2.1 : Adaptation somatosensorimotrice et codage du but des actions orofaciales

Résumé/Abstract	102
2.1.1 Introduction	103
Codage du but de l'action	103
Adaptation et processus de contrôle prédictif	108
Cadre de l'étude	108
2.1.2 Méthode	109
Protocole	109
Analyses statistiques	110
2.1.3 Résultats	111
Réseaux neuronaux des mouvements orofaciaux	112
Somatotopie sensorimotrice	114
Effet de Répétition Suppression	116
2.1.4 Discussion	118
Actions orofaciales	118
Répétition Suppression	118
Limites de l'étude	120
2.1.5 Conclusion	121

Chapitre 2.2 : Contrôle sensorimoteur adaptatif et but des actions vocaliques ...

Résumé/Abstract	122
2.2.1 Introduction	123
Codage du but des actions vocaliques lors de la perception et de la production de parole	124
Cadre de l'étude	126
2.2.2 Méthode	126
Protocole	126

Analyses statistiques	129
2.2.3 Résultats	131
Perception et production des voyelles	131
Effet de Répétition Suppression	136
2.2.4 Discussion	139
Réseaux des systèmes de perception et de production	139
Représentations différenciées entre voyelles	140
Répétition Suppression	141
Limites de l'étude	146
2.1.5 Conclusion	146

3^{ème} Partie : Rôle fonctionnel de la voie dorsale lors de la perception

de la parole	147
---------------------------	-----

Avant-propos	149
--------------------	-----

Chapitre 3.1 : Implication de la voie dorsale dans les mécanismes

d'adaptation sélective de la parole	150
--	-----

Résumé/Abstract	150
3.1.1 Introduction	151
Cadre de l'étude	153
3.1.2 Méthode	154
Protocole	154
Acquisition des données	157
Analyse des données	159
3.1.3 Résultats	161
Tâche de perception catégorielle	161
Tâche de décision syllabique	161
3.1.4 Discussion	162
3.1.5 Conclusion	165

Chapitre 3.2 : Plasticité et implication du système moteur orofacial

lors de la perception de la parole	166
---	-----

Résumé/Abstract	166
3.2.1 Introduction	167
Cadre de l'étude	168
3.1.2 Méthode	168
Protocole	168
Analyses statistiques	171
3.2.3 Résultats	173
Scores perceptifs	173
Temps de réaction	173
3.2.4 Discussion	173
3.2.5 Conclusion	174

Discussion & Perspectives	177
--	-----

Réseaux du contrôle moteur orofacial, de la production et de la perception de voyelles .	180
Rappel des résultats	180
Limites et Perspectives	182

Somatotopie et phonémotopie	184
Rappel des résultats	184
Limites et Perspectives	185
Boucles correctives et codage des buts sensorimoteurs	186
Rappel des résultats	187
Limites et Perspectives	189
Implication causale de la voie dorsale lors de la perception de la parole	190
Rappel des résultats	191
Limites et Perspectives	191
Conclusion	192
Bibliographie	193
Liste des Publications	215

Liste des figures

Figure I.1 : Non-linéarité quantique entre paramètres articulatoires et acoustiques	5
Figure I.2 : Mécanismes de coarticulation et patterns formantiques de syllabes synthétiques.....	7
Figure I.3 : Représentation schématisée des patterns acoustiques pour la perception duplex	7
Figure I.4 : Architecture simplifiée de la théorie de la perception pour le contrôle de l'action.....	10
Figure I.5 : Système auditif périphérique et central	13
Figure I.6 : Système de neurones miroirs chez le singe macaque	15
Figure I.7 : Effet de résonance motrice lors de la perception de parole.....	17
Figure I.8 : Activations somatotopiques des régions motrices lors de la perception de parole	17
Figure I.9 : Somatotopie et rôle causal des régions motrices lors de la perception de parole	18
Figure I.10 : Modèle de double voie de perception de la parole d'Hickok et Poeppel.....	19
Figure I.11 : Modèle de double voie de Rauschecker et Scott	21
Figure I.12 : Modèle d'analyse-par-synthèse de Skipper et al.	21
Figure I.13 : Schéma de correction du mouvement et mise à jour du modèle inverse	23
Figure I.14 : Musculature orofaciale et voie pyramidale de contrôle moteur orofacial	28
Figure I.15 : Modèle DIVA de production de la parole.....	31
Figure I.16 : Modèle de prédiction en série de Tian et Poeppel	32
Figure I.17 : Modèle génératif de contrôle moteur de Price, Crinion et Mac Sweeney	33
Figure I.18 : Modèle hiérarchique de contrôle moteur par feedback de Hickok	34
Figure 1.1 : Représentation de l'ordre des acquisitions lors d'une session IRM	41
Figure 1.2 : Paradigme de sparse sampling	43
Figure 1.3 : Etapes de prétraitements des images fonctionnelles	44
Figure 1.4 : Système de référence anatomique et paramètres de mouvements	45
Figure 1.5 : Régions corticales et sous-corticales impliquées dans le contrôle moteur orofacial	48
Figure 1.6 : Organisation somatotopique sensorimotrice dorso-ventrale orofaciale.....	49
Figure 1.7 : Effet principal de la tâche motrice	56
Figure 1.8 : Réseaux cérébraux des mouvements orofaciaux	57
Figure 1.9 : Somatotopie sensorimotrice (analyse par régions d'intérêts).....	60
Figure 1.10 : Effet principal du délai d'acquisition.....	61
Figure 1.11 : Topographie neurale observée lors de la perception auditive de voyelles	70
Figure 1.12 : Espace acoustique F1/F2 des voyelles préenregistrées	73
Figure 1.13 : Activations spécifiques et communes lors de la perception et production des voyelles.....	78
Figure 1.14 : Activations distinctes lors de la perception et production des voyelles.....	80
Figure 1.15 : Activations lors des tâches de perception et de production pour les voyelles fermées, mi-fermées et mi-ouvertes.....	82
Figure 1.16 : Régions cérébrales activées lors des tâches de perception et de production pour les voyelles antérieures non arrondies, antérieures arrondies et postérieures arrondies	82
Figure 1.18 : Régions cérébrales activées lors de la tâche de perception selon les 3 délais d'acquisition de 4s, 5s et 6s	83
Figure 1.17 : Activations pour les 9 voyelles lors des tâches de perception et de production.....	84

Figure 1.19 : Régions cérébrales activées lors de la tâche de production selon les 3 délais d'acquisition de 4s, 5s et 6s	86
Figure 1.20 : Régions cérébrales différemment activées lors de la tâche de perception et de production en fonction des délais d'acquisition de 4s, 5s et 6s.....	86
Figure 2.1 : Stimulations pariétales et prémotrices induisant l'envie ou l'exécution de mouvements	104
Figure 2.2 : Modèles explicatifs de l'effet de 'répétition-suppression'	106
Figure 2.3 : Effet RS lors de l'observation et de l'exécution de mouvements manuels.....	107
Figure 2.4 : Régions cérébrales activées lors de l'exécution de mouvements orofaciaux.....	112
Figure 2.5 : Effet principal de la tâche motrice	114
Figure 2.6 : Somatotopie sensorimotrice (analyse par régions d'intérêts).....	115
Figure 2.7 : Régions cérébrales activées et effet RS lors de la répétition des gestes orofaciaux	117
Figure 2.8 : Espace acoustique F1/F2 des voyelles préenregistrées	127
Figure 2.9 : Activations spécifiques et communes lors de la perception et production des voyelles	132
Figure 2.10 : Effet principal des traits d'ouverture et d'antériorité pour les voyelles perçues et produites	135
Figure 2.11 : Effet RS lors de la perception et de la production des voyelles.....	137
Figure 3.1 : Plan expérimental de l'étude TMS	155
Figure 3.2 : Pourcentage moyen des réponses /ba/ en fonction des 11 stimuli du continuum.....	156
Figure 3.3 : Représentations des régions cérébrales stimulées.....	158
Figure 3.4 : Dispositif TMS.....	159
Figure 3.5 : Réponses /ba/ et temps de réaction dans la tâche de décision syllabique	160
Figure 3.6 : Effet d'adaptation sélective observé lors de la tâche de décision syllabique	160
Figure 3.7 : Activités EMG lors de l'entraînement moteur de la langue et des lèvres.....	170
Figure 3.8 : Activités EMG moyennes des muscles labiaux et linguaux lors des entraînements.....	171
Figure 3.9 : Résultats comportementaux de la tâche de décision syllabique	172
Figure D.1 : Activations lors des tâches de mouvements orofaciaux et de production de voyelles.	181
Figure D.2 : Activations lors des tâches de perception et de production de voyelles	182
Figure D.3 : Somatotopie sensorimotrice des articulateurs supralaryngés et vocaliques	184
Figure D.4 : Effet RS lors des tâches de mouvements orofaciaux et de production de voyelles	187
Figure D.5 : Effet RS lors des tâches de perception et de production de voyelles.....	188

Liste des tables

Table 1.1 : Coordonnées des pics d'activation observés pour les quatre tâches motrices orofaciales et communs à ces tâches (analyse de conjonction).....	58
Table 1.2 : Somatotopie sensorimotrice orofaciale. Valeurs moyennes des centres de gravité des pics d'activation maximale	59
Table 1.3 : Effet principal du délai d'acquisition	61
Table 1.4 : Valeurs acoustiques moyennes des voyelles préenregistrées	73
Table 1.5 : Pics d'activation observés pour les tâches de perception et de production des voyelles et communs à ces deux tâches (analyse de conjonction)	79
Table 1.6 : Coordonnées des régions différenciellement activées lors des tâches de perception et de production des voyelles	81
Table 1.7 : Coordonnées des régions différenciellement activées lors de la tâche de production en fonction du degré d'ouverture	81
Table 1.8 : Coordonnées des pics d'activation observés pour les tâches de perception et de production en regard de trois délais d'acquisition.....	85
Table 2.1 : Coordonnées des pics d'activation maximale communs aux actions orofaciales	113
Table 2.2 : Coordonnées des pics d'activation montrant un effet RS pour les gestes orofaciaux	116
Table 2.3 : Valeurs acoustiques moyennes des voyelles préenregistrées	127
Table 2.4 : Coordonnées des pics d'activation maximale communs aux tâches de perception et de production des voyelles	134
Table 2.5 : Coordonnées des régions différenciellement activées lors de tâches de perception et de production des voyelles en fonction des degrés d'ouverture et d'arrondissement/antériorité	135
Table 2.6 : Coordonnées des régions présentant un effet de RS lors de la perception et production des voyelles	139

- Introduction -

Introduction

Les premières étapes d'acquisition de la parole sont traditionnellement considérées comme dépendantes d'une co-structuration des représentations motrices et sensorielles - auditives, visuelles et somatosensorielles - des unités de parole. Cette co-structuration s'établirait de manière progressive chez l'enfant au travers de mécanismes imitatifs et d'apprentissage sensorimoteur. A partir des premières capacités motrices et de discrimination acoustique et phonétique, ces premières étapes exploratoires permettraient une maturation et spécialisation des systèmes perceptifs et moteurs, par l'établissement d'associations entre buts moteurs et sensoriels des unités de parole perçues et produites. La maturation du système phonologique de l'enfant découlerait ainsi de l'imprégnation linguistique ambiante et d'une spécification progressive et conjointe des actes moteurs et des cibles sensorielles pertinentes.

Néanmoins, la persistance d'un tel couplage fonctionnel entre systèmes sensoriels et moteur chez l'adulte reste à ce jour débattue. En effet, les systèmes de perception et de production de la parole ont été par le passé largement étudiés indépendamment l'un de l'autre, et le plus souvent en dehors de tout cadre conversationnel et communicatif pourtant fonctions premières de l'acte de parole. Du point de vue du fonctionnement cérébral, les premières théories neurologiques, établies il y a plus d'un siècle autour de l'étude de patients aphasiques et appuyant l'implication possible de modules corticaux spécialisés respectivement dans les processus de perception et de production de la parole, ont sans doute influencé l'idée d'une séparation fonctionnelle entre ces deux systèmes.

L'hypothèse principale défendue dans la présente thèse est, au contraire, celle, chez l'adulte, d'une coordination nécessaire des représentations sensorielles et motrices aussi bien pour la spécification, la compréhension et la production des unités de parole. En appui de cette hypothèse, de nombreux modèles phonétiques, psycholinguistiques et neurobiologiques récents de la perception et/ou de la production de la parole appuient l'idée d'un couplage fonctionnel entre systèmes sensoriels et moteur et ont pour point commun l'importance attribuée aux interactions sensorimotrices. Dans ces modèles, la production de parole impliquerait la mise en œuvre de mécanismes de contrôle moteur en ligne permettant d'évaluer les conséquences sensorielles prédites des actions planifiées en les comparant avec les retours sensoriels effectifs et, de là, de corriger et réguler si nécessaire la réalisation des unités produites. Lors de la perception de la parole, des mécanismes inférentiels de simulation motrice permettraient de déterminer, ou tout du moins de contraindre, l'interprétation phonétique des entrées sensorielles.

Dans ce cadre théorique, l'objectif premier de nos travaux était de déterminer un possible couplage fonctionnel des systèmes de perception et de production des voyelles du Français, par l'examen de l'instanciation et du codage des représentations de parole au sein du système nerveux central et plus particulièrement des systèmes sensoriels et moteurs.

Avant de décrire les études réalisées dans le cadre de cette thèse, dont les bases théoriques et les hypothèses sous-jacentes seront détaillées dans chacune des sous-parties du présent ouvrage, nous proposons dans ce préambule un bref compte-rendu des études et modèles phonétiques, psycholinguistiques et neurobiologiques ayant trait à la question de la nature sensorimotrice des systèmes de perception et de production de la parole. Un guide de lecture explicitant les différents chapitres et sous-chapitres de la présente thèse clôturera cette partie introductive.

La perception de la parole sous l'angle des interactions sensorimotrices

Un problème fondamental et ancien en parole concerne la difficulté à décrire de manière simple les relations existantes entre sons et phonèmes. Du fait de la grande variabilité du signal acoustique de parole, les processus de décodage acoustico-phonétique à partir des propriétés et informations issues du signal acoustique restent en effet à ce jour encore largement incompris. Cette variabilité du signal acoustique de parole provient de nombreuses causes, parfois sous-jacentes et imbriquées. En premier lieu, la variabilité provient évidemment des différences anatomiques et de stratégies articulatoires entre locuteurs (chaque conduit vocal est unique et les stratégies articulatoires liées à la réalisation d'un énoncé linguistique diffèrent parfois grandement). De plus, pour un même locuteur, la production de tout message linguistique implique des mécanismes de co-articulation entre phonèmes produits, soit une combinaison de gestes articulatoires à débit temporel élevé et largement parallélisé. Cette co-articulation consiste en l'enchaînement fluide de phonèmes qui ne sont pas produits de manière parfaitement sérielle. Afin d'optimiser l'effort d'articulation, le passage articulatoire entre deux phonèmes voisins va être lissé, leur production emboîtée dans le temps, "co-articulée". Par ailleurs, ce phénomène de "lissage" implique également un phénomène de "réduction articulatoire" qui consiste en le fait de basculer dans l'articulation du phonème suivant, sans donc atteindre la cible articulatoire du son de parole précédent, ce pour assurer un passage fluide. Si ces mécanismes de co-articulation permettent de minimiser l'effort articulatoire produit par le locuteur, cette combinaison de gestes articulatoires est transformée de manière fortement non linéaire en une séquence de sons dans laquelle la correspondance entre sons et gestes est dès lors peu transparente pour l'auditeur... et pour le chercheur (voir par exemple, Perkell et Klatt, 1986).

Approches auditives, théories motrices et sensorimotrices

Cette variabilité inhérente du signal acoustique de parole et les relations complexes existant entre sons et phonèmes ont entraîné ces cinquantes dernières années de nombreux débats, parfois vifs, quant à la nature auditive et/ou motrice des processus mis en œuvre dans le décodage des informations phonétiques à partir des entrées sensorielles. Ainsi, certains modèles phonétiques et psycholinguistiques de la perception de la parole proposent une invariance contenue intrinsèquement dans le signal acoustique que le cerveau de l'auditeur détecte en reconstituant des valeurs acoustiques porteuses de sens phonétique (par exemple, Diehl et Kluender, 1989; Massaro, 1998; Kuhl, 2000; pour une revue, voir Diehl, Lotto et Holt, 2004). A l'opposé, pour d'autres chercheurs, l'interprétation phonétique des entrées sensorielles de la parole est déterminée, ou tout du moins partiellement contrainte, par les connaissances procédurales articulatoires de l'auditeur (par exemple, Liberman et al., 1967; Liberman et Mattingly 1985; Fowler, 1986, 1996; Liberman et Whalen, 2000; Schwartz et al., 2002, 2012; pour des revues récentes voir Perrier, 2005; Galantucci, Fowler et Turvey, 2006; Schwartz, Sato et Fadiga, 2008, 2011). Nous proposons dans le présent sous-chapitre une brève présentation de ces différentes approches et théories ainsi qu'une discussion des limites des hypothèses posées par chacune.

Les approches générales auditives

Pour les tenants des **approches générales auditives**¹ (par exemple, Diehl et Kluender, 1989; Massaro, 1998; Kuhl, 2000; Diehl, Lotto et Holt, 2004), l'objet de la perception de la parole serait de nature

¹ Les auteurs des diverses "approches générales auditives", en proposant un cadre général d'un fonctionnement perceptif général, ne proposent pas de modèles de la perception. Leur objectif n'est pas de fournir des modèles concrets explicatifs de la perception de la parole, mais plutôt de s'opposer à certains points théoriques des théories motrices (voir prochaine

purement auditive. Les invariants phonémiques seraient récupérés et directement accessibles dans le signal acoustique et traités uniquement par le système auditif, les processus de décodage acoustico-phonétique ne faisant donc pas appel à des mécanismes moteurs articulatoires, ni inférentiels, ni directs².

De plus, ces mécanismes de décodage acoustico-phonétique reposeraient sur des mécanismes auditifs et d'apprentissage perceptifs non propres aux seuls sons de parole mais liés à des prédispositions et capacités perceptives plus générales³. Ce postulat théorique se fonde notamment sur deux éléments empiriques fondamentaux liés à la perception catégorielle de la parole, soit la capacité à mieux discriminer à écart acoustique constant deux phonèmes situés entre deux catégories phonologiques plutôt qu'à l'intérieur de celles-ci (Liberman et al., 1957) :

- Premièrement, le fait que la perception catégorielle n'est pas un phénomène spécifique à la parole. En effet, divers travaux remontant aux années 70 (Stevens et Klatt, 1974; Pisoni, 1977) ont montré que des sons non langagiers peuvent également être perçus de manière catégorielle.
- Et de plus, que la perception catégorielle est également présente chez des espèces animales, y compris pour des sons de parole (Kuhl et Miller, 1975, 1978).

Pris ensemble, ces deux classes d'arguments suggèrent ainsi que la perception catégorielle de la parole se base sur des principes perceptifs généraux et non spécifiques ni à l'homme ni à la parole. Puisque la perception catégorielle de la parole avait été initialement proposée par les tenants des approches motrices comme une preuve d'un mode spécial de la perception de la parole (voir ci-dessous), ces études ont de plus largement contribué à réévaluer, voire même rejeter, ce postulat théorique⁴ (voir Galantucci, Fowler et Turvey, 2006 pour une revue critique).

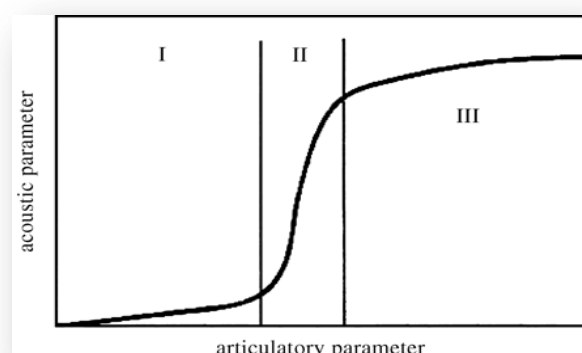


Figure I.1 : Schéma de non-linéarité quantique dans l'appariement entre paramètres articulatoires du conduit vocal et sortie acoustique. Les régions I et III sont stables du point de vue acoustique, par rapport aux modifications dans la dimension articulatoire. La région II est en revanche acoustiquement instable et de ce fait est supposée représenter une frontière catégorielle entre les deux catégories phonétiques issues des régions I et III (figure empruntée à Stevens, 1989).

Un autre argument indirect permettant d'expliquer la récupération des invariants phonémiques dans le signal de parole provient de l'examen des systèmes phonétiques des langues du monde. En effet,

section). Ainsi, puisque ces positions ne contiennent pas suffisamment de propositions explicatives pour être falsifiables, elles ne se qualifient pas de *théories* et sont par conséquent considérées plutôt comme des *approches* (voir Diehl, Lotto et Holt, 2004).

² De manière particulière, Massaro (1989; voir aussi Massaro et Chen, 2008) propose un modèle de la perception de la parole ("fuzzy logical model of perception") dans lequel la perception catégorielle de la parole est expliquée en termes de reconnaissance probabiliste de patterns prototypiques visuels et acoustiques dont l'intégration permet l'identification de l'unité de parole produite. Cela suppose la mémorisation des sons de parole en terme de prototypes d'unités de parole, au sein desquels des traits acoustiques (ou visuels) peuvent varier dans une certaine mesure, mais dont l'ensemble des valeurs déterminent leur place au sein d'une catégorie.

³ Un cas particulier concerne toutefois la théorie de l'invariance acoustique (Stevens et Blumstein, 1978; Blumstein et Stevens, 1979) qui considère la perception de la parole comme reposant sur les propriétés supposées invariantes du signal acoustique mais ne dénie pas un rôle possible du système moteur articulatoire quand l'information acoustique est ambiguë ou incomplète.

⁴ Ainsi, Liberman et collègues (1972) écrivaient avant ces études: "Presumably, they [animals] lack the special processor necessary to decode the speech signal. If so, the perception of speech must be different from ours. They should not hear categorically, for instance, and they should not hear the [di]-[du] patterns... as two segment syllables which have the first segment in common".

malgré les possibilités infinies d'articulation du conduit vocal, on retrouve dans la plupart des langues du monde certains phonèmes récurrents et les explications proposées mettent en avant les propriétés auditives – et non motrices – de ces unités. Ainsi, pour expliquer ce phénomène, Stevens (1972, 1989) a introduit dans la théorie quantique de la perception de la parole la notion de 'quanta' par rapport aux relations non-linéaires entre la configuration du conduit vocal et les sorties acoustiques qui en résultent. A l'intérieur de certaines zones articulatoires stables, des déplacements importants de certains articulateurs (par exemple, l'ouverture de la mandibule) n'ont pas de conséquences acoustiques importantes (régions I et III sur la Figure I.1), tandis que dans d'autres configurations, des déplacements articulatoires moindres auront des incidences acoustiques considérables (région II sur la Figure I.1). Pour Stevens (1972, 1989), ces zones de stabilité et d'instabilité acoustique auraient permis l'établissement de frontières catégorielles phonétiques naturelles, les zones de stabilité perceptive permettant une discrimination naturelle entre catégories phonétiques. De même, en ce qui concerne les systèmes vocaliques des langues du monde (voir par exemple, Vallée, 1994; Schwartz et al., 1997a, 1997b pour un examen de l'organisation et typologie des systèmes vocaliques), Liljencrants et Lindblom (1972) ont proposé une explication à la quasi-universalité des voyelles [i]-[a]-[u] liée à leurs propriétés perceptives extrêmement différentes. Ils ont ainsi postulé dans la théorie de la dispersion adaptative que l'existence de ces 3 voyelles au sein des langues du monde pouvait être expliquée par des raisons pragmatiques de maximisation des distances perceptives entre voyelles et donc d'optimisation de leur discrimination perceptive. La théorie quantique et la théorie de la dispersion mettent ainsi au niveau de la distinctivité acoustico-perceptive la base de l'existence d'universaux dans les systèmes sonores des langues du monde.

Les théories motrices de la parole

A l'opposé des approches générales auditives, certains chercheurs ont proposé que l'objet de la perception de la parole est de nature articulatoire (Liberman et al., 1962, 1967; Liberman et Mattingly, 1985; Fowler, 1986, 1996; Liberman et Whalen, 2000). Ici, les invariants phonémiques seraient récupérés grâce à un appariement entre les gestes articulatoires du locuteur à l'origine du signal de parole perçu et les connaissances et représentations motrices activées par ce signal chez l'auditeur.

La théorie motrice de la perception de la parole(Liberman et al., 1962, 1967; Liberman et Mattingly, 1985; Liberman et Whalen, 2000) suppose que les informations issues du signal de parole sont mises en correspondance avec les représentations motrices des gestes articulatoires à l'origine du signal acoustique perçu (en termes de commandes neuromotrices globales plus qu'en termes de gestes articulatoires précis, permettant de minimiser partiellement les problèmes de variabilité articulatoire; voir Liberman et Mattingly, 1985). En d'autres termes, face à la non-linéarité entre propriétés acoustiques du signal de parole et représentations phonétiques sous-jacentes, Liberman et collègues proposent de considérer les invariants non pas en termes acoustiques mais en termes moteurs. Leur hypothèse est donc que la perception de la parole repose sur un appariement des représentations phonétiques et des commandes motrices neuronales et sur une récupération des gestes articulatoires du locuteur par l'auditeur.

La communication est ainsi rendue possible (1) grâce à un mécanisme de parité des représentations phonétiques et motrices invariantes entre locuteur et auditeur (pour reprendre les termes de Liberman et Whalen, 2000 : "Ce qui compte pour le locuteur doit compter pour l'auditeur"), (2) permettant que les mêmes représentations (de nature motrice dans cette théorie) soient utilisées aussi bien en production qu'en perception de la parole.

Différents arguments empiriques ont été à l'origine de la théorie motrice de la perception de la parole. Un premier argument essentiel concerne les mécanismes de co-articulation et le problème de la non-linéarité entre signal acoustique et phonème.

Prenons pour exemple un contexte syllabique simple consonne-voyelle /di/ et /du/. Les propriétés acoustiques relatives au phonème /d/ diffèrent en fonction de la voyelle suivante /i/ ou /u/ (voir

Figure 1.2). L'exemple des syllabes /di/ et /du/ montre que les valeurs formantiques du phonème /d/ varient selon le contexte syllabique, alors que la commande neuromotrice pour ce phonème est, elle, similaire, ayant pour but dans les deux cas de positionner l'apex de la langue vers les dents supérieures (lieu d'articulation apico-dental). Par conséquent, pour les auteurs, le problème de co-articulation peut être dépassé dès lors que l'on considère que l'objet de perception récupéré par l'auditeur n'est pas une représentation auditive (qui semble variable) mais une commande neuromotrice (qui serait, elle, invariante) nécessaire à la réalisation du phonème perçu. Les auteurs proposent donc un lien très fort entre système de perception et de production de la parole, qui partageraient les représentations des unités de parole en termes de commandes neuromotrices.

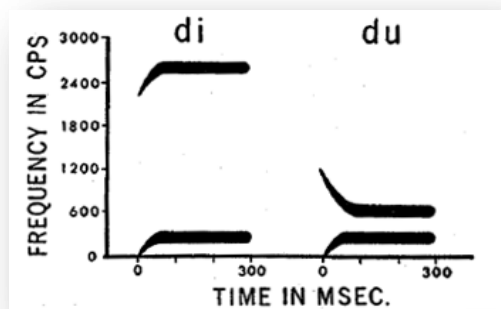


Figure 1.2: Représentation schématique des patterns formantiques F_1/F_2 de deux syllabes synthétiques /di/ et /du/ conçues par Liberman et collaborateurs (1967). De manière à simuler un mécanisme de co-articulation, les valeurs formantiques F_2 montante ou descendante de l'attaque consonantique (/d/) sont déterminées par la voyelle suivante (/i/ ou /u/). Malgré des différences acoustiques des deux syllabes, les sujets perçoivent le même phonème initial /d/. Cet exemple a nourri l'idée des théories motrices que l'objet de la perception de la parole soit de nature motrice (Figure empruntée à Liberman et al., 1967).

Un second postulat de la théorie motrice de la perception de la parole, apparu dans une version modifiée de cette théorie (Liberman et Mattingly, 1985), est que la perception de la parole soit fondamentalement différente de la perception d'autres types de sons, que cette capacité soit innée et spécifique aux humains qui seuls auraient développé un module unique et spécifique dédié à la parole⁵.

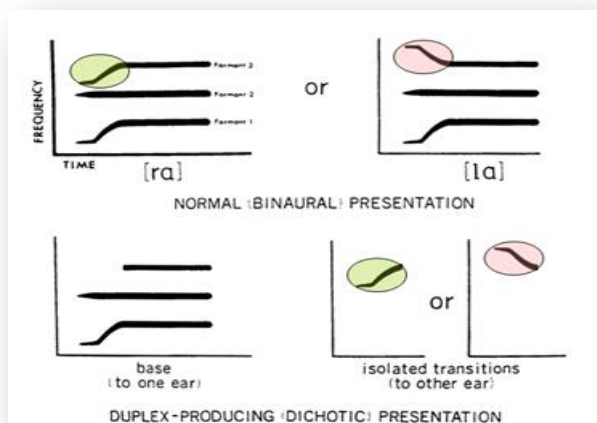


Figure 1.3: Représentation schématique des patterns acoustiques pour la perception duplex du [ra] (en vert) et du [la] (en rouge). L'oreille A reçoit la base syllabique avec des transitions formantiques de deux formants (en bas à gauche) qui est phonologiquement ambiguë, et l'oreille B reçoit la troisième transition (en bas à droite), qui en soi correspond à un sifflement mais permet de désambigüiser le signal perçu par l'oreille A. L'auditeur perçoit à la fois le signal de parole reconstitué mais aussi le sifflement. Ainsi, il s'agit d'une double perception de ce sifflement, qui est perçu une fois en tant qu'élément permettant de compléter la base syllabique et une deuxième fois en tant que sifflement en soi (Figure adaptée de Liberman, Isenberg et Rakerd, 1981).

⁵ Ce postulat est à rattacher à la théorie de la modularité de l'esprit émise par Fodor (1983), considérant l'existence de modules spécialisés dans certaines fonctions cognitives, largement innés et fonctionnant de manière autonome et cloisonnée les uns des autres. De manière importante, l'aspect "spécial" de la parole tel que proposé par Liberman et collègues a été vivement débattu et de nombreuses études ont contribué à réévaluer, voire abandonner, ce postulat théorique (voir Galantucci, Fowler et Turvey, 2006 pour une revue critique). Citons à titre d'exemple contre le caractère inné de la perception de la parole, les études sur les enfants sauvages montrant que la capacité à développer un langage nécessite un entraînement intense en bas âge et l'existence d'une période critique pour cet apprentissage (voir Malson, 1964). De même, pour l'aspect spécifiquement humain, comme décrit précédemment, il a été montré que certains animaux perçoivent la parole humaine de manière catégorielle (par exemple, Kuhl et Miller, 1975, 1978; Lotto, Kluender et Holt, 1997). Enfin, il est également possible de citer ici l'hypothèse de Stanislas Dehaene sur le recyclage neuronal et l'évolution de structures cérébrales présentes chez les primates non humains vers des fonctions cognitives langagières et arithmétiques (Dehaene, 2004).

Cette proposition vient notamment de la découverte de l'effet "duplex" (Rand, 1974), qui consiste en une perception double d'un signal acoustique décomposé. Dans un exemple classique de cet effet, on présente de manière dichotique (c'est-à-dire en distribuant l'information entre les deux oreilles) deux signaux acoustiques issus d'un même signal de parole: une oreille A reçoit la base syllabique avec des transitions formantiques de deux formants, la troisième transition, qui désambigüe le signal, étant fournie à la seconde oreille B (voir Figure 1.3). De cette manière, acoustiquement, l'oreille A reçoit un signal acoustique phonologiquement ambigu, alors que l'oreille B reçoit un signal acoustique qui en soi correspond à un sifflement. Au lieu de fusionner simplement ces deux signaux, l'auditeur crée un double percept (la perception duplex) et perçoit dans l'oreille A le signal de parole reconstitué mais également dans l'oreille B un sifflement. Cette perception duplex, soit le fait qu'un même fragment du signal acoustique de parole soit interprété à la fois comme unité de sens phonétique et comme un sifflement, a été utilisée comme argument empirique en faveur de l'existence de deux modules, l'un spécialisé dans le traitement phonétique des sons de parole, l'autre relié à des mécanismes auditifs généraux.

Porter et Castellanos (1980) et Porter et Lubker (1980) ont également apporté des éléments en faveur de la théorie motrice de la perception de la parole en réalisant des tâches de « shadowing », où les participants devaient répéter le plus rapidement possible des séquences de parole perçues, sans attendre la fin du signal du parole ni tenir compte de son interprétation. Ces auteurs ont montré que le délai entre le début de la présentation du signal et le début de l'articulation était extrêmement court (inférieur à 200ms). Ce résultat a été pris comme argument en faveur d'une récupération rapide des gestes articulatoires à l'origine du signal perçu (voir même directe, dans l'hypothèse où les commandes motrices seraient déjà présentes dans le signal acoustique, voir ci-dessous).

Issue des mêmes laboratoires Haskins qu'Alvin Liberman, Carol Fowler (1986, 1996) a proposé une théorie motrice alternative, **la théorie directe réaliste de la perception de la parole**. En désaccord sur le caractère spécifique de la parole et sur la nature intrinsèque de l'objet perceptif, elle a proposé que les processus à la base de la perception des sons de parole et des sons non-langagiers reposent sur les mêmes mécanismes perceptifs généraux, également liés à d'autres modalités sensorielles comme la vue ou le toucher. Pour Fowler, la perception de la parole ne repose donc pas sur un module spécifique mais sur des prédispositions et mécanismes perceptifs plus généraux, et consiste en la reconnaissance directe de la nature de la source de la perturbation ressentie/perçue dans l'environnement (par exemple, la lumière pour la vision, la peau pour le toucher et l'air pour l'audition). En conséquence, plutôt que de récupérer les commandes neuronales reliées aux gestes articulatoires du locuteur comme dans la théorie motrice de la perception de la parole, Fowler propose une reconnaissance directe de la structure de l'effecteur à l'origine des perturbations de l'environnement, que ce soit la modification de la forme du conduit vocal pour la parole ou un autre médium qui modifie le signal acoustique perçu dans le cas d'un son non-langagier⁶. En d'autres termes, dans cette théorie, ce sont les gestes articulatoires d'un locuteur lors de la réalisation d'un message linguistique qui structurent le signal acoustique, cette information étant supposée contenue dans le signal acoustique et directement perceptible par l'auditeur sans recours à des mécanismes inférentiels. Cette reconnaissance requiert donc une connaissance des propriétés fonctionnelles de l'objet perçu et une connaissance préalable des actions potentielles en lien avec l'objet. Cette hypothèse apparaît cohérente avec la proposition de Viviani et Stucchi (1992) pour qui la perception est guidée par les connaissances procédurales motrices de l'observateur. Ces auteurs ont en effet montré que la connaissance implicite de contraintes biomécaniques, notamment les règles cinématiques et les configurations articulatoires sous-jacentes aux mouvements du corps, biaise

⁶ Cette idée est inspirée de James Gibson (1966, 1979) et son "approche écologique de la perception visuelle" qui se base sur la notion de constance perceptive et peut s'expliquer uniquement dans la relation de celui qui perçoit avec son environnement. Il a notamment proposé le concept d'"affordances", c'est-à-dire l'ensemble des possibilités d'action d'un environnement qui sont directement et objectivement perçues par celui qui observe. Ainsi, la perception est abordée par la fonctionnalité, par le potentiel d'agir sur l'environnement.

notre perception des mouvements. Enfin, outre le rejet de mécanismes inférentiels dans la récupération des gestes articulatoires du locuteur, Fowler s'est également séparée de la théorie motrice classique de Liberman et collègues par rapport à un possible statut modulaire et spécifique de la parole. Pour appuyer leur hypothèse, Fowler et Rosenblum (1990) ont ainsi démontré que l'effet de la perception duplex existe également pour des signaux acoustiques non langagiers (par exemple, un claquement de la porte), ce qui démontre la non-spécificité de cet effet à la seule perception de la parole.

La Théorie de la Perception pour le Contrôle de l'Action (PACT)

Nous l'avons vu, à la base de la théorie motrice de Liberman et collègues est la non-linéarité entre signal acoustique et phonème, liée notamment à l'existence de mécanismes de co-articulation. Pour les auteurs, le fait de percevoir un même phonème malgré des différences acoustiques évidentes (dans différents contextes co-articulés, voir Figure 1.2) est contraire à la possibilité que l'information phonétique soit entièrement contenue dans le signal acoustique de parole. De là, l'hypothèse de mécanismes inférentiels permettant de récupérer les gestes articulatoires, ou plus précisément les commandes neuromotrices nécessaires à la réalisation du phonème perçu. Néanmoins, il est à noter que la théorie motrice déplace mais ne résout pas totalement ce problème d'invariance entre sons et phonèmes. En effet, dans certaines situations de production perturbée, il a été montré que des stratégies articulatoires différentes entre sujets sont mises en œuvre pour la réalisation d'un même phonème (voir par exemple, Savariaux, Perrier et Orliaguet, 1995; Savariaux et al., 1999). Plus important encore, certaines configurations articulatoires différentes permettent d'obtenir le même son de parole (voir Mielke, Baker et Archangeli, sous presse; décrit dans Schwartz et al., 2012). Pour certains phonèmes, il n'est donc pas possible de pouvoir récupérer univoquement une seule configuration des gestes articulatoires et du conduit vocal.

D'autre part, le postulat d'équivalence perceptive et motrice dans un contexte de co-articulation peut également être interprété à la lumière des approches générales auditives de la perception de la parole. Suivant Brunswik (1956) et sa proposition d'expliquer la reconnaissance des objets visuels malgré la grande variabilité des images visuelles, la constance perceptive d'un objet pourrait résulter d'une recombinaison d'attributs multiples. Lors de la perception de parole, il a ainsi été proposé que l'auditeur, grâce à des mécanismes d'apprentissage perceptifs, puisse maintenir de manière similaire une constance perceptive phonémique des objets perçus, pourtant variables sur le plan acoustique. Ces mécanismes d'apprentissage perceptifs pourraient également expliquer pourquoi différentes espèces d'oiseaux, des chinchillas ou des macaques (Kuhl et Miller, 1978; Kuhl et Padden, 1982; Kluender et al., 1987; Dooling et al., 1995) sont capables d'apprendre à catégoriser les sons de parole suite à un entraînement intensif (par exemple, le voisement ou le lieu d'articulation des consonnes occlusives /b/, /d/ et /g/) et ce, sans être pourvu ni de système articulatoire de production de parole ni de module spécialisé.

Ainsi, mis à part le postulat de Liberman et collègues sur l'existence d'un module spécialisé et inné dans la perception de la parole, postulat contredit par de nombreuses études, de nombreux débats continuent d'opposer les tenants des approches générales auditives et des théories motrices quant à la nature des processus mis en œuvre dans les processus de décodage acoustico-phonétique (voir par exemple, Diehl, Lotto et Holt, 2004; Galantucci, Fowler et Turvey, 2006; Schwartz, Sato et Fadiga, 2008, 2011).

Face aux approches générales auditives et aux théories motrices, la **théorie de la perception pour le contrôle de l'action** (PACT, Perception for Action Control Theory; Schwartz et al., 2002, 2012) propose une position de compromis quant à la nature des processus mis en œuvre lors de la perception de la parole, processus ni purement auditifs, ni purement moteurs mais bien sensorimoteurs. Cette théorie se base notamment sur l'existence de connaissances procédurales à la fois motrices et auditives, partagées par le locuteur et l'auditeur et issues de mécanismes d'apprentissage et d'une co-structuration des systèmes et représentations motrices et sensorielles

notamment lors de l'acquisition de la parole. De manière importante dans cette théorie, si les processus de décodage acoustico-phonétique exploitent les capacités du système auditif, l'interprétation phonétique finale des entrées sensorielles serait en partie contrainte par les connaissances procédurales motrices de l'auditeur (Schwartz et al., 2002, 2012; Schwartz, Sato et Fadiga, 2008). En ce sens, cette théorie peut être qualifiée de théorie constructiviste⁷ de perception de la parole en proposant que les représentations articulatoires de l'auditeur contraignent partiellement l'interprétation des entrées sensorielles à travers des mécanismes de simulation motrice interne.

En d'autres termes, pour la théorie de la perception pour le contrôle de l'action, la perception de la parole reposerait sur des mécanismes multisensoriels (auditifs et visuels⁸) et serait contrainte par des connaissances procédurales motrices sous-jacentes. Ainsi, "les représentations de parole seraient intrinsèquement sensorimotrices, ni purs produits sensoriels, ni purs objets moteurs inférés, mais des percepts multimodaux régularisés par l'action" (Sato, 2004).

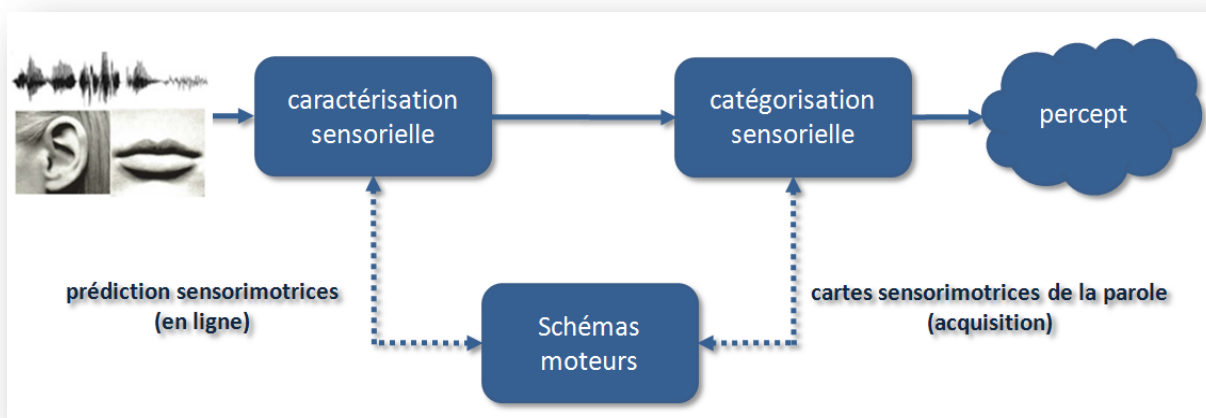


Figure I.4 : Architecture simplifiée de la perception de la parole selon la théorie de la perception pour le contrôle de l'action (Figure adaptée de Schwartz et al., 2012).

La théorie de la perception pour le contrôle de l'action est notamment basée sur différents postulats théoriques (voir Figure I.4):

- Premièrement, une co-structuration des systèmes sensoriels et moteurs de la parole les premières années de vie, au travers de mécanismes imitatifs et d'apprentissage sensorimoteur et en regard des contraintes phonologiques environnementales. Cette co-structuration a pour effet la mise en place de cartes sensorimotrices reliant représentations sensorielles et motrices de la parole.
- Deuxièmement, la perception structure l'action. En d'autres termes, le système de perception fournit au système de production des gabarits auditifs, les gestes de parole sont donc façonnés et sélectionnés dans les systèmes phonologiques en regard de leurs propriétés perceptives auditives et visuelles et leur valeur fonctionnelle communicative⁹.

⁷ La notion de constructivisme découle notamment de la théorie de l'apprentissage développée par Piaget et autres psychologues du développement et suppose une (re)construction des connaissances du sujet au fur et à mesure des apprentissages à partir d'éléments déjà intégrés lors d'expériences passées.

⁸ En dehors du rôle du système moteur lors de la perception de la parole, les recherches portant sur les mécanismes d'intégration multisensorielle, notamment audio-visuelle, ne seront pas abordées dans la présente thèse (pour des revues, voir Möttonen, 2004; Besle, 2007; Basirat, 2010; Fort, 2011).

⁹ Un exemple donné par Schwartz et collègues (2012) est celui du triangle vocalique [i]-[a]-[u] présent dans la quasi-totalité des langues du monde. Une explication simple à cette ubiquité est que ces trois voyelles ont des propriétés acoustiques et motrices maximale distinctes permettant une catégorisation différenciée optimale. Au contraire, la triade vocalique [y]-[a]-[u] possède les mêmes caractéristiques motrices mais n'apparaît dans aucune langue. Pour Schwartz et collègues, cela s'explique par la faible distinctivité auditive de ces voyelles.

- Enfin, l'action met en forme la perception. Le système perceptif se construit en fonction des contraintes motrices des gestes orofaciaux. Si ce postulat théorique est principalement lié à la période d'acquisition de la parole chez l'enfant, même à l'âge adulte la perception peut impliquer le recours à de telles connaissances procédurales motrices permettant au locuteur d'extraire, de prédire et d'intégrer les événements sensoriels perçus en ligne de manière cohérente.

Codage cérébral de la perception de la parole

La recherche sur le cerveau s'est développée en phases successives au fur et à mesure de l'évolution des techniques disponibles. Si les premières recherches menées par dissection et coloration sélective (méthode de Golgi, par exemple) du XIX^{ème} siècle ont permis une première description anatomique du cerveau, il a fallu attendre les travaux de neurophysiologie à l'aide de micro-électrodes (donc typiquement invasifs) pour mieux comprendre le fonctionnement détaillé des neurones ou de réseaux de neurones. En parallèle, d'autres travaux, guidés par l'étude des conséquences de lésions cérébrales, ont permis d'identifier certaines régions du cerveau en rapport avec leur fonction sensorielle, motrice ou cognitive (notamment les travaux de Broca, de Wernicke et de Lichtheim sur des patients aphasiques). Plus récemment, de nouvelles techniques d'électroencéphalographie (EEG) puis d'imagerie par résonance magnétique fonctionnelle (IRMf), de magnétoencéphalographie (MEG) et de stimulation magnétique transcrânienne (TMS) ont rendu possible l'observation du fonctionnement cérébral avec une excellente résolution spatiale ou temporelle.

Appliquée à l'étude de la parole, l'utilisation de ces techniques a notamment permis de démontrer l'activité de régions spécifiques auditives, au sein du gyrus temporal supérieur (aire de Wernicke), et motrices, au sein du gyrus frontal inférieur (aire de Broca), de l'hémisphère gauche lors de la perception et lors de la production de parole. Ces résultats semblent ainsi confirmer l'implication de modules corticaux spécialisés respectivement dans les processus de perception et de production de la parole et du langage, un résultat déjà suggéré par les théories neurologiques issues de l'étude de patients aphasiques depuis plus d'un siècle. Néanmoins, l'hypothèse d'une séparation fonctionnelle stricte de ces régions lors de ces tâches est de plus en plus contestée. En ce qui concerne la perception de la parole, de nombreuses études récentes ont en effet démontré l'activation des régions frontales impliquées dans la planification et l'exécution des gestes articulatoires lors de la perception d'unités de parole.

A la lumière de ces études, ce sous-chapitre a pour objectif de présenter brièvement les bases corticales de la perception auditive de la parole, par la présentation d'études démontrant l'implication de régions corticales du système moteur lors de la perception et reconnaissance d'unités de parole puis par l'examen de modèles neurobiologiques récents postulant l'existence d'un couplage fonctionnel entre systèmes sensoriels et moteur de la parole¹⁰.

En premier lieu, nous rappellerons ici brièvement les mécanismes neuronaux basiques concernant les traitements auditifs des signaux acoustiques, qui, de manière très simplifiée, sont traités en périphérie par l'organe récepteur (la cochlée), avant d'arriver via le système nerveux périphérique (le nerf auditif) jusqu'au système nerveux central (du tronc cérébral jusqu'au cortex auditif; voir Figure 1.5; pour une revue des principes d'anatomie et de fonctionnement du système nerveux, voir Purves et al., 2005; Bear, Connor et Paradiso, 2002; Baciú, 2011).

Les premiers traitements du signal acoustique sont pris en charge par le **système auditif périphérique**. Schématiquement, les ondes sonores sont reçues par le pavillon de l'oreille externe et

¹⁰ Ce bref aperçu des traitements auditifs cérébraux ne rend pas compte de l'ensemble des traitements neuronaux liés à la perception de la parole et du langage, ni de l'ensemble des régions impliquées dans ces traitements et de leur connectivité fonctionnelle et temporelle (pour des revues détaillées, voir par exemple Zatorre, Belin et Penhune, 2002; Scott et Johnsrude, 2003; Scott et Wise, 2004; Démonet, Thierry et Cardebat, 2005; Price, Thierry et Griffiths, 2005; Chevillet, Riesenhuber et Rauschecker, 2007; Zatorre et Gandour, 2008; Zatorre, 2007; Rauschecker et Scott, 2009; Turkeltaub et Coslett, 2010; Rauschecker, 2011, 2012).

entrent par transmission aérienne le long du conduit auditif pour faire vibrer la membrane du tympan. Cette vibration mécanique est transmise à la chaîne des osselets de l'oreille moyenne¹¹ attachée au tympan (le marteau¹², l'enclume et l'étrier) qui par système de levier renforcent et transmettent ces vibrations mécaniques jusqu'à la fenêtre ovale d'entrée à l'oreille interne. Ici, deux organes se trouvent interconnectés : l'appareil vestibulaire pour l'équilibre et la cochlée pour l'audition. La cochlée est une cavité creuse qui comporte trois canaux parallèles séparés par des membranes (pour éviter le contact entre le liquide les remplissant), deux rampes (vestibulaire et tympanique) contenant la périlymphe. Le canal cochléaire rempli d'endolymphe se trouve au milieu, délimité par la membrane de Reissner et basilaire. Cette dernière supporte l'organe de Corti, qui contient les récepteurs de l'audition proprement dits, les cellules ciliées. Celles-ci sont ancrées sur la membrane basilaire et possèdent chacune une centaine de stéréocils, des excroissances attachées par leur extrémité à la membrane tectoriale. Lors d'une survenue d'une onde sonore, la membrane basilaire se déplace entraînant une déformation des cellules ciliées, car leurs stéréocils étant attachés à la membrane tectoriale figée, leurs corps cellulaires suivent le mouvement de la membrane basilaire. Cette inclinaison provoque une dépolarisation de ces cellules et c'est à ce niveau que s'effectue la transduction mécano-électrique. En effet, les cellules ciliées forment des synapses avec des neurones dont les corps cellulaires se situent en proximité dans les ganglions spiraux et dont les axones constituent les fibres qui pénètrent dans le nerf auditif par où l'information est transmise le long du tract auditif aux centres nerveux. Une propriété constante de tous les relais de la voie auditive est son organisation spécifique aux fréquences sonores, qui est la conséquence de l'organisation anatomique de la cochlée. En effet, au sein de la cochlée, enroulée en elle-même sur deux tours et demi, la membrane basilaire augmente son diamètre de la base à l'apex, tel que la base de la membrane répond optimalement à des fréquences aiguës, et l'apex à des fréquences graves. Les cellules ciliées, tapissant régulièrement la membrane basilaire, reflètent cette organisation. Les fréquences basses excitant prioritairement les cellules ciliées se trouvant le plus éloignées de la base, et les fréquences aiguës celles qui sont proches de l'apex. Cette disposition anatomique de la cochlée décide d'une organisation fonctionnellement 'tonotopique' qui est respectée par les fibres séparées du nerf auditif et se poursuit tout au long de la voie auditive.

Système auditif central. La cochlée stimule les premiers neurones de la voie auditive¹³ qui vont relayer ces informations acoustiques aux noyaux cochléaires ipsilatéraux du tronc cérébral. Là, il y a décussation (croisement des voies) de la majeure partie des fibres pour atteindre les noyaux de l'olive supérieure contralatérale et, via le colliculus inférieur, le corps genouillé médian du thalamus¹⁴ (en parallèle, les fibres restantes aboutissent aux structures ipsilatérales). Du thalamus, l'information acoustique va enfin être transmise au cortex auditif primaire, lui aussi organisé tonotopiquement, en vue des premiers traitements auditifs corticaux. Le cortex auditif qui reçoit et analyse les informations acoustiques est structuré hiérarchiquement en aires primaires, secondaires et tertiaires (ou associatives), anatomiquement organisées de façon concentrique dans les parties supérieures et moyennes du lobe temporal (voir Figure I.5).

¹¹ A noter la présence de la trompe d'Eustache, cavité descendant au niveau de l'oreille moyenne vers le pharynx, qui n'est pas impliquée directement dans le traitement auditif du son mais dans l'ajustement de la pression d'air de l'oreille moyenne par rapport au milieu ambiant.

¹² Un ajustement de l'intensité peut avoir lieu en situation de niveau sonore élevé, par le réflexe d'atténuation. Dans ce cas, il y a un rétrocontrôle sur l'oreille moyenne qui consiste à contracter à la fois le muscle tenseur du tympan (attaché d'un côté à l'os de la cavité de l'oreille moyenne et de l'autre côté au marteau) et le muscle stapédien (ancré d'un côté à la cavité de l'oreille moyenne et de l'autre côté à l'étrier) ce qui mène à la rigidification des structures de l'oreille moyenne et de ce fait diminue l'énergie transmise à l'oreille interne.

¹³ Outre la voie auditive, les mêmes informations sont dirigées en convergence avec des informations provenant d'autres systèmes sensoriels vers le système réticulaire qui fait la sélection d'informations arrivant au cortex afin d'éviter une saturation des systèmes des traitements perceptifs (due à une surstimulation) et pour maintenir l'attention ou réguler le système d'éveil/sommeil.

¹⁴ D'autres informations auditives sont envoyées au colliculus supérieur (où elles convergent avec les projections visuelles) et vers le cervelet.

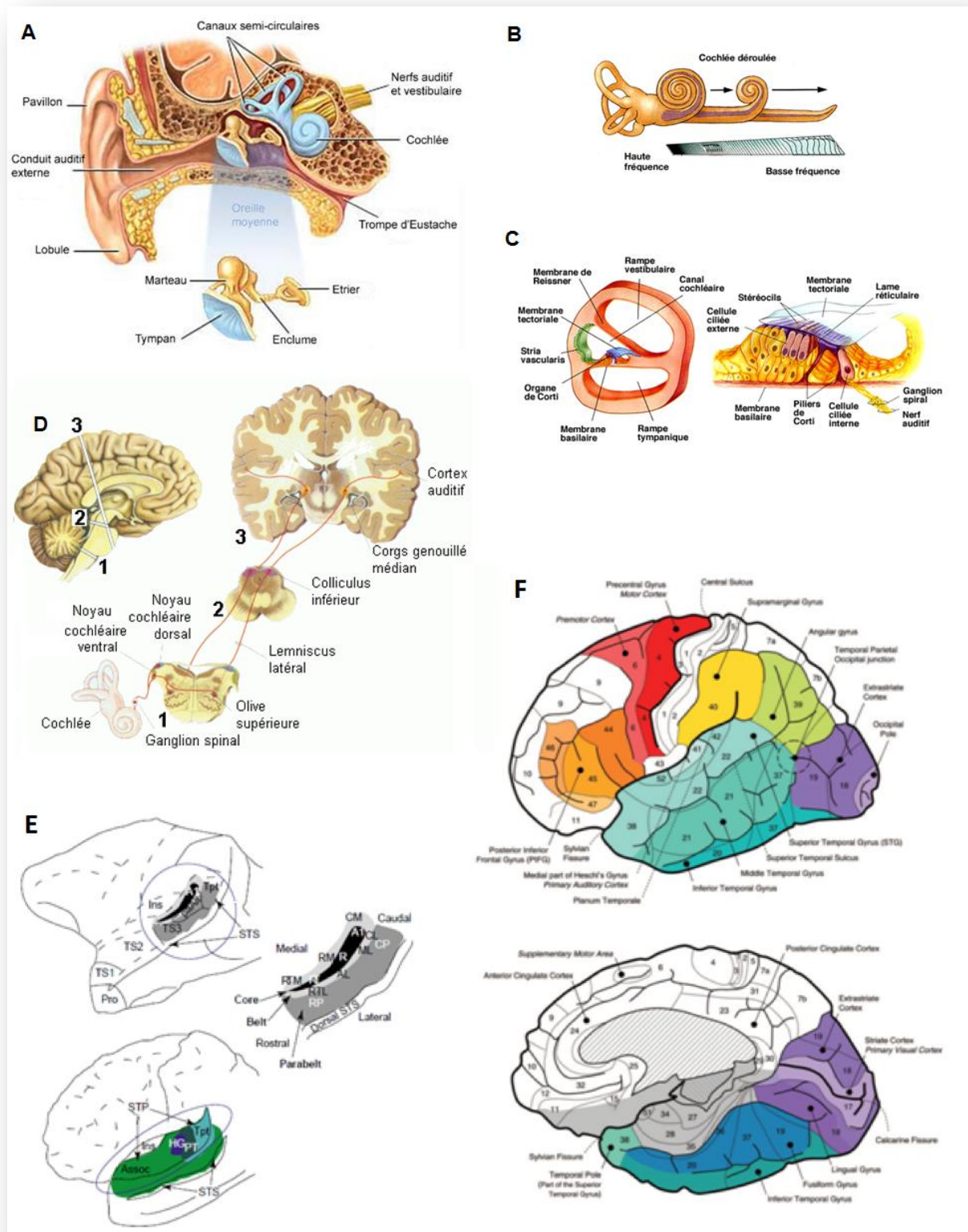


Figure 1.5 : (A) Oreille externe, moyenne et interne. (B) Cochlée et membrane basilaire déroulée avec fréquences préférentielles. (C) Coupe verticale à travers la cochlée et zoom sur le récepteur auditif, l'organe de Corti avec les cellules ciliées. (D) Voie auditive, depuis l'entrée de l'onde sonore dans le pavillon de l'oreille externe jusqu'à l'arrivée au cortex auditif. (E) Régions temporelles auditives chez le singe macaque (haut) et l'homme (bas). (F) Principales régions cérébrales impliquées dans la perception de la parole et du langage. Gauche : Vue latérale de l'hémisphère gauche. Droite : Vue sagittale médiane de l'hémisphère droit. Les nombres indiquent les aires de Brodmann (sources des figures : www.haenggeli.org/FAQs/AnatOreille.html; www.vetopsy.fr/sens/audi/audi_oreille3.php; www.unige.ch/cyberdocuments/theses2001/CrottazHerbetteS/these_body.html; Scott et Johnsrude, 2003; Démonet, Thierry et Cardebat, 2005).

A noter que les connaissances du système auditif cortical humain sont notamment inférées à partir des études sur les primates non humains, dont les aires corticales auditives sont organisées de manière concentrique (aires auditives centrales / 'core', à la ceinture / 'belt' et régions auditives entourant la ceinture / 'parabelt'; voir Scott et Johnsrude, 2003). Le premier traitement auditif cortical est effectué au niveau de l'aire auditive centrale (qui correspond à l'aire auditive primaire, l'aire BA 41¹⁵, située bilatéralement au niveau du gyrus de Heschl, encore appelé gyrus temporal transverse) qui répond préférentiellement aux sons purs. Elle est entourée par une ceinture d'aires auditives secondaires (l'aire BA 42), qui reçoivent des informations depuis l'aire primaire et le thalamus auditif, mais effectuent des traitements moins précis en terme tonotopique en favorisant des analyses spectrales des sons plus complexes. Enfin, les régions auditives associatives, entourant la ceinture (aire BA 22), effectuent des traitements de plus haut niveau des stimuli (notamment de décodage acoustico-phonétique) et impliquant des processus d'attention sélective, de mémoire auditive, d'intégration multimodale, les parties postérieures du gyrus et sulcus temporal supérieur faisant interface avec des régions associatives visuelles, somatosensorielles.

Système de perception de parole. Il est largement admis que le signal acoustique de la parole est traité par des voies parallèles spécialisées dans différents niveaux de traitements. Les premiers traitements acoustiques de bas niveau, non spécifiques aux sons de parole, auraient ainsi lieu dans les gyri temporaux supérieurs dorsaux bilatéraux (incluant le gyrus de Heschl et des régions antérieures du planum temporale). La localisation exacte des régions cérébrales spécialisées dans le décodage acoustico-phonétique et la catégorisation des phonèmes perçus reste néanmoins controversée. En effet, certains chercheurs proposent que les signaux acoustiques complexes (non spécifiques à la parole) sont d'abord traités dans la partie médio-latérale du gyrus/sulcus temporal supérieur et que les représentations phonétiques des sons de parole seraient instanciées suite à ces traitements dans la partie antérieure du STG/STS (Scott et Johnsrude, 2003; Rauschecker et Scott, 2009; Rauschecker, 2011). Au contraire, d'autres chercheurs soutiennent que les catégories phonétiques sont instanciées dans la partie postérieure du gyrus temporal moyen et au sein de la partie adjacente du sillon temporal supérieur (Belin et Zatorre, 2000; Hickok et Poeppel, 2007). Néanmoins, à ce jour, aucune étude ne permet véritablement d'appuyer ni de contredire l'existence d'une voie auditive antérieure (Scott et Johnsrude, 2003) ou postérieure (Hickok et Poeppel, 2007) impliquée dans les processus de décodage acoustico-phonétique lors de la perception de voyelles.

D'autre part, une spécialisation hémisphérique quant aux traitements de la parole et du langage a été mise en évidence (pour des revues, voir Zatorre, Belin et Penhune, 2002; Meyer, 2008; Perrone-Bertolotti, 2011) : l'hémisphère droit semble plus spécialisé dans le traitement des informations supra-segmentales, notamment prosodiques, et la distinction des hauteurs tonales (pour la perception de la musique par exemple), et l'hémisphère gauche dans les traitements spectro-temporels précis des stimuli auditifs et les processus de discrimination phonétique (voir par exemple, Zatorre et Belin, 2001; Zatorre et Gandour, 2008). Poeppel (2003) souligne, que le traitement acoustique de la parole se fait sur différents niveaux d'échelles temporelles et propose que l'implication différentielle des deux gyri temporaux peut être expliquée dans leur préférence de fenêtre d'intégration temporelle : l'hémisphère droit traiterait les informations acoustiques au sein de "longues" fenêtres de traitement (entre 150 et 250ms) alors que l'hémisphère gauche peut analyser des changements acoustiques survenant entre 20 et 40ms (voir aussi Boemio et al., 2005; Britton et al., 2009).

Implication du système moteur dans la perception de la parole

Face aux modèles phonétiques et psycholinguistiques de la perception de la parole présentés précédemment, la plupart des modèles neurobiologiques actuels s'accordent sinon sur leur rôle

¹⁵ Le terme BA, pour Brodmann area, fait référence à la terminologie utilisée dans l'atlas cytoarchitectonique de Brodmann (1909). Dans cet atlas, un numéro allant de 1 à 59 a été attribué à chaque région du cortex selon son organisation cellulaire (nombre de couches, épaisseurs des couches, arborisation dendritique...).

précis tout du moins sur l'activation et le recrutement de structures motrices corticales lors de la perception des unités de parole. Avant une présentation de ces différents modèles neurobiologiques, nous présenterons dans ce sous-chapitre différentes études ayant permis d'apporter des éléments nouveaux en faveur de l'implication du système moteur dans l'observation et la reconnaissance de l'action de manière générale, et de la parole en particulier.

Neurones miroirs. Ces deux dernières décennies, l'utilisation des nouvelles techniques de neurophysiologie et d'imagerie et les découvertes sur le cerveau des primates humains et non humains ont fourni des arguments empiriques importants en faveur du rôle possible du système moteur dans la reconnaissance des actions. Dans ce cadre, une découverte fondamentale est celle des fameux neurones miroirs au sein des cortex prémoteur et pariétal des singes macaques, formant un système spécialisé dans l'observation et l'exécution des actions ("execution-observation matching system"; pour des revues, voir Rizzolatti, Fogassi et Gallese, 2001; Rizzolatti et Craighero, 2004; Rizzolatti et Sinigaglia, 2010).

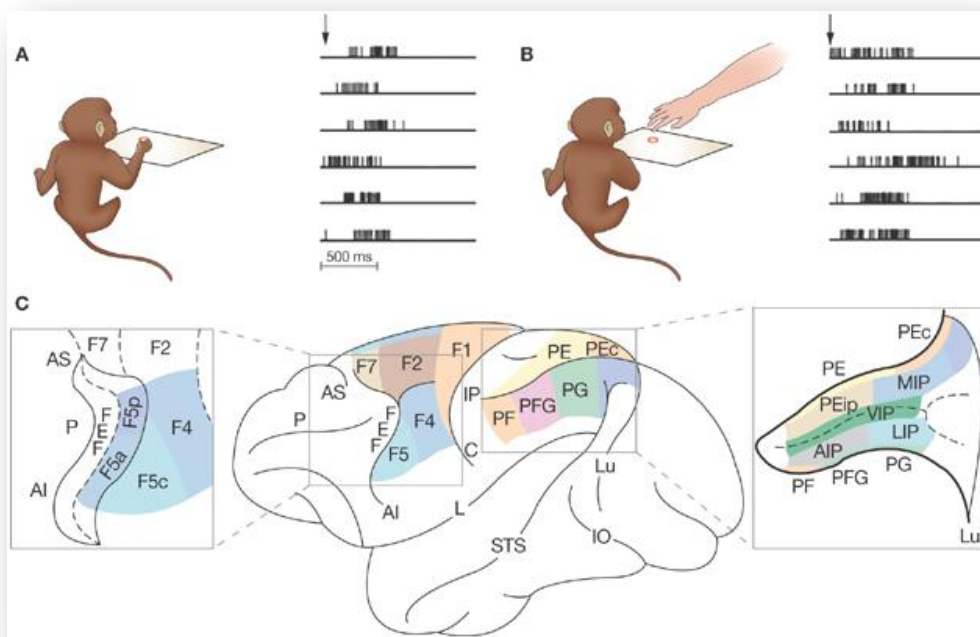


Figure 1.6 : A et B. Exemple d'un neurone miroir de l'aire F5 du singe macaque qui décharge à la fois lorsque le singe exécute une action de saisie et lorsqu'il observe un autre individu effectuer cette action. C. Vue latérale gauche du cerveau du singe macaque et des régions prémotrices et pariétales dont les régions F5 et PF/PFG contenant les neurones miroirs (Figure adaptée de Rizzolatti et al., 2008).

Rappelons que le système cortical moteur est composé d'une constellation d'aires différentes (voir Figure 1.6) distinctes par leurs propriétés fonctionnelles et leurs connexions avec les autres aires corticales et sous-corticales. De ce point de vue on peut grossièrement distinguer les aires motrices postérieures pariéto-dépendantes (F1 à F5), recevant les informations sensorielles issues du cortex pariétal et impliquées dans des processus de transformations sensorimotrices lors de la réalisation d'une action et les aires antérieures préfronto-dépendantes (F6-F7), recevant des informations du cortex préfrontal et jouant un rôle de contrôle moteur, par exemple les processus d'initiation et d'inhibition des actions. Une première découverte fondamentale antérieure à celle des neurones miroirs est que l'aire F5¹⁶ chez le singe macaque contient un répertoire des représentations d'actions manuelles et orofaciales (Rizzolatti et al., 1988). En effet, la majorité des neurones de F5 déchargent de manière spécifique non en fonction de mouvements mais durant la réalisation d'une action dirigée vers un but. Par exemple, certains neurones déchargent spécifiquement lors d'une action de

¹⁶ L'aire F5a est généralement considérée comme l'analogue de la partie opératoire de l'aire de Broca (voir Nelissen et al., 2005; Petrides, Cadoret et Mackey, 2005).

saisie d'un objet et ce avec la main gauche, la main droite ou encore la bouche, ce qui suggère que ces neurones sont capables de généraliser le but de l'action indépendamment de l'effecteur impliqué. En fait, la spécificité de réponse de ces neurones suggère l'idée que F5 constitue une sorte de vocabulaire, de lexique, des actes moteurs, certains neurones codant la catégorie générale à laquelle appartient une action donnée (par exemple la saisie ou le déplacement d'un objet), d'autres la manière de réaliser cette action (par exemple une action de saisie nécessitant une préhension de l'objet précise ou plus large; "precision grip" vs "power grip"), et enfin, d'autres étant concernés par les aspects temporels de la réalisation de cette action (par exemple, lors de la seule initiation de l'action ou durant la réalisation entière de l'action).

Une seconde découverte fondamentale antérieure à celle des neurones miroirs est l'existence de neurones sensorimoteurs, appelés neurones canoniques, qui déchargent non seulement lors de la réalisation d'une action mais également lors de la simple présentation d'un objet (Rizzolatti et al., 1988; Gentilucci et al., 1988). Il existe de plus une congruence stricte entre les deux types de réponses, motrices et visuelles, du neurone. Ainsi, les neurones activés lors de la réalisation d'une action nécessitant une configuration manuelle précise des doigts, par exemple lors de la saisie d'un anneau, le seront également lors de la simple présentation de cet objet. Alors que l'activation de ces neurones lors de la saisie ou observation d'un objet nécessitant une configuration manuelle différente, par exemple un cylindre ou une sphère, sera inexistante ou moindre. Ceci suggère donc un lien étroit entre les caractéristiques fonctionnelles d'un objet et le type d'actions liées à cet objet, ce qui va dans le sens du concept d'affordances proposé par Gibson (1966, 1979) et son "approche écologique de la perception visuelle".

Les neurones miroir sont des neurones polymodaux visuo-moteurs ou audio-visuo-moteurs localisés dans les cortex prémoteur ventral et pariétal (aires F5 et PF/PFG) du singe macaque et dont les réponses/décharges neuronales interviennent aussi bien lorsque le singe exécute des actions manuelles ou orofaciales que lorsqu'il perçoit visuellement ou entend des actions similaires exécutées par un autre individu (Di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 1996; Ferrari et al., 2003; Kohler et al., 2002; Keyser et al., 2003; Fogassi et al., 2005)¹⁷. L'existence des neurones miroirs suggère ainsi que l'observation d'actions implique partiellement les mêmes circuits neuronaux que ceux utilisés pour la réalisation de ces actions. Depuis cette découverte, de nombreuses études en neurophysiologie et neuroimagerie ont permis de mettre en évidence l'existence d'un système de neurones miroirs chez l'homme (pour des revues, voir Rizzolatti, Fogassi et Gallese, 2001; Rizzolatti et Craighero, 2004; Rizzolatti et Sinigaglia, 2010). En plus de la reconnaissance de l'action, il a été proposé que le système de neurones miroir chez l'homme jouerait un rôle fondamental dans le traitement de la parole en fournissant un mécanisme neurophysiologique à la base du principe de "parité motrice" entre individus communicants (Rizzolatti et Arbib, 1998; Rizzolatti et Craighero, 2004; Arbib, 2005; Gentilucci et Corballis, 2006; voir aussi Aboitiz et Garcia, 1997).

Activations motrices en perception. Des arguments plus directs en faveur d'un lien entre la perception et la production de la parole proviennent de l'observation de l'activation des régions motrices (la partie postérieure, operculaire, du gyrus frontal inférieur, le cortex prémoteur ventral et le cortex moteur primaire) et des régions proprioceptives liées aux mouvements orofaciaux (cortex somatosensoriel) lors de tâches de perception de la parole selon des modalités auditive, visuelle ou audio-visuelle (par exemple, Nishitani et Hari, 2002; Calvert et Campbell, 2003; Callan et al., 2003; Paulesu et al., 2003; Callan et al., 2004; Möttonen et al., 2004; Wilson et al., 2004; Ojanen et al., 2005; Pekkola et al., 2005; Skipper, Nusbaum et Small, 2005; Pulvermüller et al., 2006; Wilson et Iacoboni, 2006; Skipper et al., 2007; Callan et al., 2010; Tremblay et Small, 2011).

¹⁷ Une étude a permis de démontrer l'existence de neurones miroirs codant des actions orofaciales communicatives (Ferrari et al., 2003). Ce sont d'ailleurs à notre connaissance les seuls neurones miroirs découverts chez le singe macaque codant des actions intransitives, soit non dirigées vers un objet. En ce sens, ce résultat apparaît important par rapport au rôle possible des neurones miroirs dans l'évolution/apparition de la parole et du langage.

De manière importante, des études IRMf montrent une activité accrue des régions du système moteur inversement fonction du degré d'intelligibilité de la parole perçue lors de parole masquée ou bruitée (Binder et al. 2004; Zekveld et al., 2006). De plus, d'autres études démontrent une plus forte implication du système moteur lors de l'identification de phonèmes non natifs vs natifs (Callan et al., 2004; Wilson et Iacoboni, 2006). Enfin, d'autres études ont également montré une augmentation d'activité des régions motrices frontales lors de la perception audio-visuelle de la parole lorsque les deux modalités étaient phonétiquement conflictuelles (par exemple, Ojanen et al., 2005; Skipper et al., 2007). Face à ces résultats, le système moteur serait ainsi plus fortement recruté en condition d'écoute difficile ou en cas d'ambiguïté perceptive, dans le but de permettre une meilleure reconnaissance/interprétation phonétique des sons de parole perçus.

La technique de stimulation magnétique transcrânienne à impulsion unique (single pulse TMS) a également permis de mettre en évidence un mécanisme de "résonance motrice" (voir Figure I.7), correspondant à une augmentation de l'excitabilité des neurones du cortex moteur primaire reliés aux muscles orofaciaux labiaux et linguaux lors de la perception de la parole (Sundara, Namasivayam et Chen, 2001; Fadiga et al., 2002; Watkins, Strafella et Paus, 2003; Watkins et Paus, 2004; Roy et al., 2008; Sato et al., 2010; pour une revue de l'application de la technique de TMS lors de l'observation d'action, voir Fadiga, Craighero et Olivier, 2005). De plus, Watkins et Paus (2004), en utilisant conjointement les techniques de TMS et de tomographie par émission de positons (PET) ont montré que cette excitabilité du cortex moteur primaire labial lors de la perception de la parole était corrélée avec l'activité du gyrus frontal inférieur postérieur.

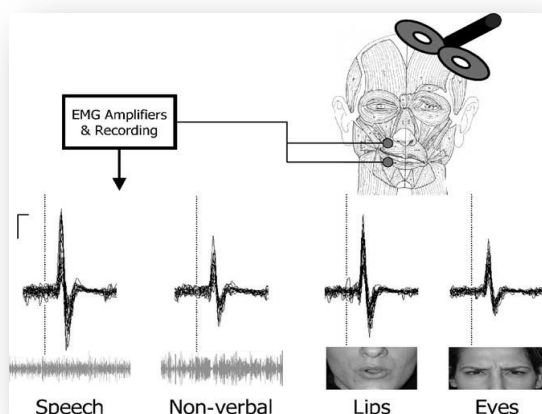


Figure I.7 : Potentiels évoqués moteurs labiaux lors d'une impulsion unique de stimulation magnétique transcrânienne sur le cortex moteur primaire contrôlant les lèvres, synchrones avec la présentation auditive ou visuelle de stimuli langagiers et non langagiers. Un effet de résonance motrice correspondant à une augmentation des potentiels évoqués moteurs est observé pour les stimuli de parole (Figure empruntée à Watkins, Strafella et Paus, 2003).

Les résultats de certaines études IRMf et par stimulation magnétique transcrânienne à impulsion unique suggèrent de plus que l'activité au sein des cortex prémoteur ventral et moteur primaire orofacial serait organisée de manière somatotopique en relation avec les effecteurs orofaciaux impliqués dans la production des sons et/ou gestes de parole perçus auditivement et/ou visuellement (Fadiga et al., 2002; Pulvermüller et al., 2006; Skipper et al., 2007; Roy et al., 2008; Sato et al., 2010; voir Figure I.8).

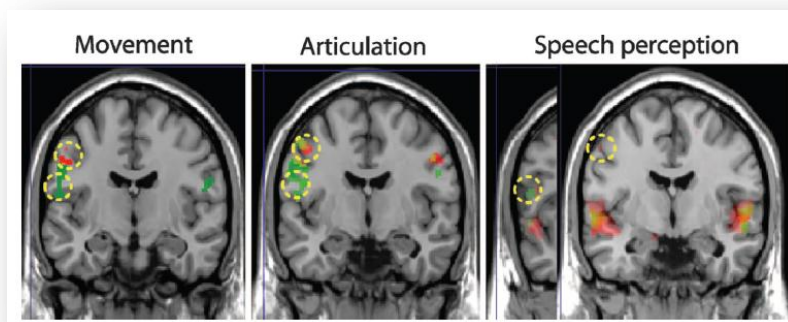


Figure I.8 : Somatotopie des activités motrices observées lors de la réalisation de mouvements articulaires labiaux (en rouge) ou linguaux (en vert), lors de la production de syllabes incluant une attaque consonantique bilabiale (en rouge) ou apico-dentale (en vert) et lors de la perception auditive de ces mêmes syllabes (Figure empruntée à Pulvermüller et al., 2006).

Rôle causal du système moteur. Si les données de neuroimagerie démontrent clairement le recrutement du système moteur lors de la perception de la parole, ces résultats sont par nature corrélationnels et ne permettent pas de démontrer un rôle médiateur, fonctionnel, du système moteur dans la reconnaissance et compréhension de la parole. Certaines études récentes utilisant les techniques de stimulation électrocorticale au cours d'opérations neurochirurgicales chez des patients éveillés, de TMS répétitive ou à double impulsion (rTMS/double pulse TMS), ont permis de démontrer qu'une modulation temporaire de l'activité neuronale de ces régions perturbe les capacités du sujet dans des tâches phonologiques "complexes" nécessitant un recrutement important de processus de segmentation et de mémoire de travail verbale (Boatmann, 2004; Nixon et al., 2004; Romero et al., 2006; Sato et al., 2009). De manière importante, des effets d'interférence lors de tâches plus simples d'identification ou de discrimination de syllabes sont également observés mais seulement dans le cas de stimuli auditifs ambigus, par exemple présentés dans du bruit blanc (Meister et al., 2007; d'Ausilio et al., 2009; Möttönen et Watkins, 2009; d'Ausilio et al., 2011, 2012).

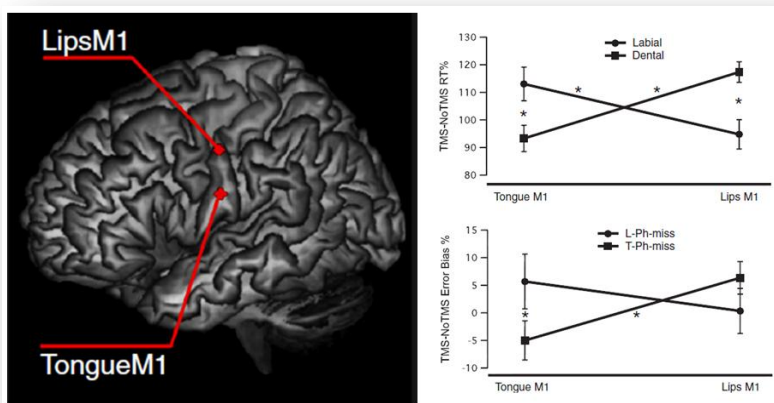


Figure 1.9 : Somatotopie et rôle causal des régions motrices primaires contrôlant les lèvres et la langue lors de la perception auditive de syllabes présentant une attaque consonantique labiale ou apico-dentale (Figure empruntée à d'Ausilio et al., 2009).

Ainsi, Meister et al. (2007) ont démontré que, par rapport à une condition contrôle sans stimulation, la stimulation répétée (900 pulses délivrés à 1Hz) du cortex prémoteur ventral gauche entraînait une diminution du pourcentage de réponses correctes (d'environ 80% à 70%) lors d'une tâche auditive subséquente de catégorisation syllabique (/pa/, /ta/, /ka/), où les syllabes étaient présentées conjointement à du bruit blanc. De même, d'Ausilio et al. (2009) ont montré une modulation spécifique des performances des sujets dans une tâche de catégorisation auditive de syllabes présentées dans du bruit blanc (/pœ/, /bœ/, /tœ/, /dœ/), suite à une stimulation en ligne par double impulsion du cortex moteur primaire gauche contrôlant soit les lèvres, soit la langue (voir Figure 1.7). Lors de l'application de la double stimulation du cortex moteur primaire lié aux lèvres, 150 et 50ms avant l'attaque consonantique, des temps de réaction et un nombre d'erreurs moindres ont été observés pour les syllabes présentant une attaque consonantique bilabiale (/pœ/, /bœ/) par rapport aux syllabes présentant une attaque consonantique apico-dentale (/tœ/, /dœ/). De plus, la stimulation du cortex moteur primaire lié au contrôle lingual a produit le résultat inverse. En accord avec les études présentées ci-dessus, cette double dissociation des résultats observés en fonction du type de syllabe perçue et de la région stimulée démontre ainsi le rôle causal du cortex moteur primaire gauche au cours de la perception de parole bruitée en relation avec son organisation somatotopique. En appui de ce résultat, Möttönen et Watkins (2009) ont également montré qu'une perturbation temporaire (900 pulses délivrés à 1Hz) de l'aire motrice primaire codant les mouvements labiaux entraînait une baisse de performances lors de la perception catégorielle de syllabes impliquant prioritairement les lèvres pour leur production (stimuli issus de continuums /ba/-/da/ ou /pa/-/ta/) mais non pour des syllabes impliquant prioritairement la langue (stimuli issus de continuums /ka/-/ga/ et /da/-/ga/). Enfin, utilisant le même protocole de double stimulation en ligne qu'utilisé dans leur précédente étude (d'Ausilio et al., 2009), d'Ausilio et collègues (2012) ont montré que la stimulation de l'aire prémotrice gauche contrôlant le larynx entraînait une diminution des

temps de réactions lors d'une tâche de similarité entre deux mêmes voyelles (/a/) mais dont la fréquence fondamentale était différente ($\pm 3\text{Hz}$)¹⁸.

Il est à noter que les modulations de performance observées dans ces études, bien que significatives, restent néanmoins limitées (de l'ordre de 10%) et que d'autres études TMS ne montrent pas de tels effets d'interférence ou de facilitation dans le cas de stimuli de parole non ambigus (Sato et al., 2009; d'Ausilio et al., 2011). Pris ensemble, les résultats de ces études démontrent néanmoins clairement le rôle fonctionnel causal des régions motrices et prémotrices de l'hémisphère gauche dans la perception de stimuli ambigus de parole et ce, conformément avec l'organisation somatotopique des régions du cortex moteur primaire.

Modèles neurobiologiques de la perception de la parole

En accord avec les études démontrant l'activation de régions du système moteur orofacial lors de la perception de la parole, la plupart des modèles neurobiologiques récents de perception de la parole postulent l'existence d'une connexion forte entre les systèmes de perception et de production. Ces modèles ont pour point commun de supposer que les liens entre ces systèmes dérivent de liens perception-action observés pour un large ensemble d'actions non linguistiques et, pour certains, l'existence de mécanismes inférentiels de simulation motrice permettant de contraindre l'interprétation phonétique des entrées sensorielles (Hickok et Poeppel, 2000, 2004, 2007; Scott et Johnsrude, 2003; Callan et al., 2004; Scott et Wise, 2004; Wilson et Iacoboni, 2006; Skipper et al., 2007; Rauschecker et Scott, 2009; Rauschecker, 2011; Hickok, Houde et Rong, 2011). Nous présenterons dans ce sous-chapitre trois de ces modèles dans lesquels le rôle du système moteur diffère sensiblement d'un modèle à l'autre.

Modèle de double voie ventrale/dorsale de Hickok et Poeppel (2000, 2004, 2007). Inspiré directement du modèle de deux voies ventrales et dorsales en vision¹⁹, le modèle de double voie de la perception auditive de la parole et du langage d'Hickok et Poeppel (2000, 2004, 2007) postule l'existence de deux voies de traitement, respectivement lexico-sémantique ventrale et d'intégration sensorimotrice dorsale (voir Figure I.10).

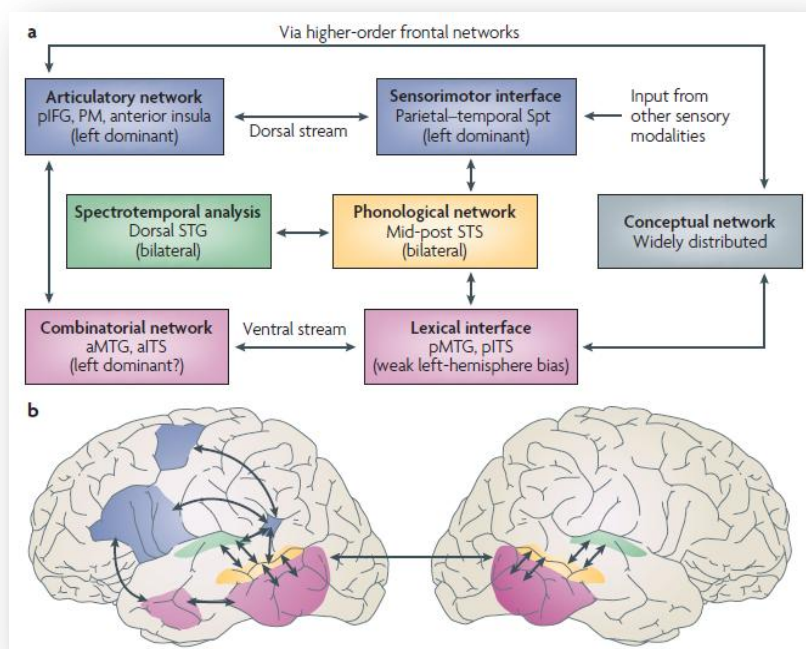


Figure I.10 : Schéma du modèle de double voie de perception de la parole et du langage d'Hickok et Poeppel (Figure empruntée à Hickok et Poeppel, 2007).

¹⁸ A noter que les auteurs ont utilisé ici pour la stimulation de l'aire prémotrice du larynx les coordonnées observées dans l'étude de Brown, Ngan et Liotti, (2008). Pour une discussion sur la vraisemblance de ces coordonnées, voir le Chapitre 1.1.

¹⁹ Initialement proposé par Ungerleider et Mishkin (1982) et repris ensuite par Goodale et Milner (1992).

Dans ce modèle, les premiers traitements de décodage du signal perçu consistent en une analyse spectrotemporelle du signal acoustique au sein du cortex auditif primaire et secondaire dans la partie dorsale moyenne du gyrus temporal supérieur bilatéral (vert). Les traitements phonologiques et de décodage acoustico-phonétique sont alors mis en œuvre dans la partie postérieure des sulci temporaux supérieurs (jaune). A partir de là, deux voies de traitement, lexico-sémantique ventrale et d'intégration sensorimotrice dorsale, sont postulées :

- la voie ventrale (rose) apparie les informations auditives/phonologiques avec les représentations conceptuelles lexico-sémantiques localisées au niveau des parties postérieures du gyrus temporal moyen et du sulcus temporal inférieur et des parties antérieures de ces mêmes régions principalement dans l'hémisphère gauche.
- la voie dorsale (bleu) apparie les représentations auditives/phonologiques et articulatoires au niveau des régions prémotrices et frontales inférieures de l'hémisphère gauche, la jonction temporo-pariétale gauche (ou aire SPT pour 'sylvian-parietal-temporal') servant d'interface sensorimotrice.

Pour Hickok et Poeppel (2007), la voie dorsale jouerait un rôle critique dans l'appariement des représentations auditives et motrices lors de l'acquisition de la parole et du langage. Chez l'adulte cependant, cette voie dorsale n'est pas considérée comme une composante déterminante de la perception de la parole. Pour les auteurs, les mécanismes d'intégration sensorimotrice mis en œuvre au sein de cette voie joueraient un rôle dans des tâches méta-phonologiques "non-écologiques"²⁰, soit lorsque l'auditeur doit explicitement utiliser des processus moteurs pour maintenir actives des représentations auditives, par exemple dans des tâches phonologiques ou de mémoire de travail verbale (Démonet et al., 1992, 1994; Zatorre et al., 1992; Paulesu, Frith et Frackowiak, 1993; Burton et al., 2000 – pour une revue, voir Poldrack et al., 1999; Démonet, Thierry et Cardebat, 2005; Vigneau et al., 2006). Enfin, un autre rôle de cette boucle sensorimotrice concerne la production de la parole et, plus spécifiquement, des mécanismes de correction et de régulation articulatoires (Hickok, Houde et Rong, 2011; Hickok, 2012).

Modèle à double voie antérieure/postérieure de Scott, Rauschecker et collègues (2003, 2009, 2011). Basé sur des études liées aux traitements auditifs et visuels chez le primate non-humain, un modèle de double voie de la perception auditive de la parole a également été proposé par Scott, Rauschecker et collègues (Scott et Johnsrude, 2003; Scott et Wise, 2004; Rauschecker et Scott, 2009; Rauschecker, 2011).

Ce modèle propose l'existence d'une voie antérieure (ou voie 'quoi'/'what') impliquée dans l'identification des objets auditifs et dans la compréhension de la parole intelligible, alors qu'une voie postéro-dorsale (ou voie 'où'/'where') serait impliquée dans la localisation des sources auditives. De plus, comme pour la voie dorsale d'Hickok et Poeppel (2007), cette voie postéro-dorsale serait également impliquée dans les mécanismes d'intégration entre représentations auditives et motrices notamment lorsque le signal acoustique est dégradé, lorsque les phonèmes appartiennent à un système phonologique non natif ou lorsque les mots entendus sont rares et peu fréquents.

De manière plus spécifique, la voie antérieure implique, suite à des traitements acoustiques de bas niveau dans le gyrus de Heschl, des processus de décodage acoustico-phonétique dans une région antérieure du gyrus supérieur temporal et dans des régions frontales inférieures supposées contenir des représentations invariantes des catégories phonétiques. Ces informations seraient ensuite envoyées au cortex prémoteur ventral pour être appariées avec les représentations articulatoires correspondantes. Les conséquences sensorielles de l'exécution simulée de telles commandes motrices seraient alors envoyées au lobule pariétal inférieur par le biais de copies d'efférence. En direction inverse, une voie dorsale impliquerait le lobule pariétal inférieur en tant que région

²⁰ Hickok et Poeppel considèrent en effet que les tâches méta-phonologiques, par exemple d'identification phonétique ou syllabique, ne sont pas naturelles et que leur utilisation ne permet pas véritablement d'étudier les processus mis en œuvre lors de la perception de la parole.

d'intégration sensorimotrice, où des comparaisons entre les copies d'efférence prédictives provenant du cortex prémoteur ventral et les informations sensorielles réelles auraient lieu et permettraient une désambiguïsation de l'information phonétique.

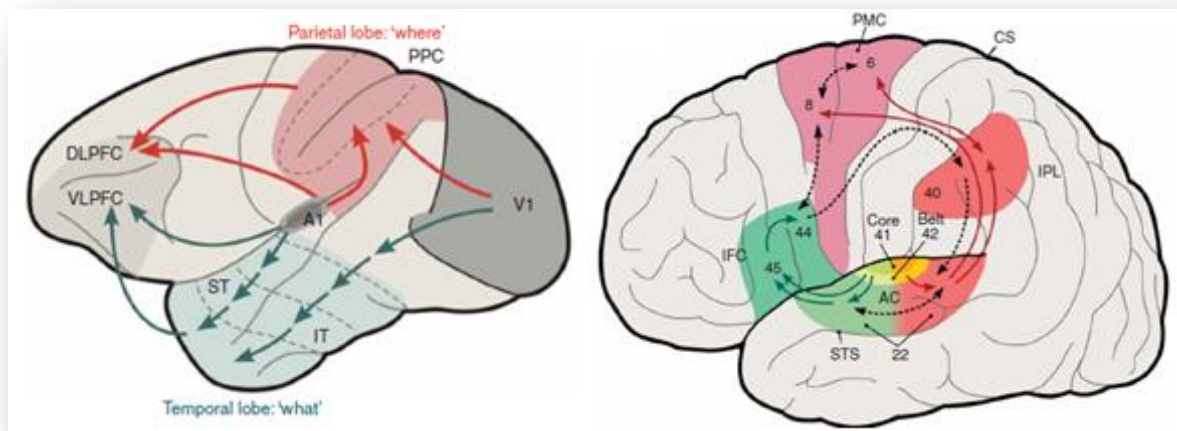


Figure I.11 : Gauche : double voies visuelles et auditives chez le singe. La voie antérieure 'quoi' (en bleu) permet l'identification de l'objet visuel ou sonore. La voie postérieure 'comment' (en rouge) permet de localiser cette source et d'appréhender les caractéristiques spatiales de l'objet perçu. Droite : Modèle de double voie de Rauschecker et Scott (2009) appliqué à la perception auditive de la parole chez l'homme.

Il est enfin à noter que Scott et collègues proposent un rôle crucial du système moteur et plus largement de la voie postéro-dorsale dans le traitement de la parole lors des tours de parole pendant les conversations, donc dans des situations naturelles de langage (Scott, McGettigan et Eisner, 2009). Dans ce cas, le système moteur permettrait de passer de manière fluide de l'écoute à la production, mais aussi de coordonner d'autres actions périphériques, de manière consciente ou non : il s'agit de coordonner la respiration, de faire converger la prononciation entre locuteurs. De manière importante, les auteurs soulignent que l'implication de cette double voie et notamment du système moteur ne se limite pas à la parole, mais peut survenir dans d'autres actions communicatives et vocales, comme le rire ou l'expression des émotions et le chant.

Modèle d'analyse-par-synthèse de Skipper et collaborateurs (2007). Dans le modèle d'analyse-par-synthèse proposé par Skipper et al., (2007), l'entrée auditive est tout d'abord traitée par le cortex auditif qui élabore des hypothèses phonémiques.

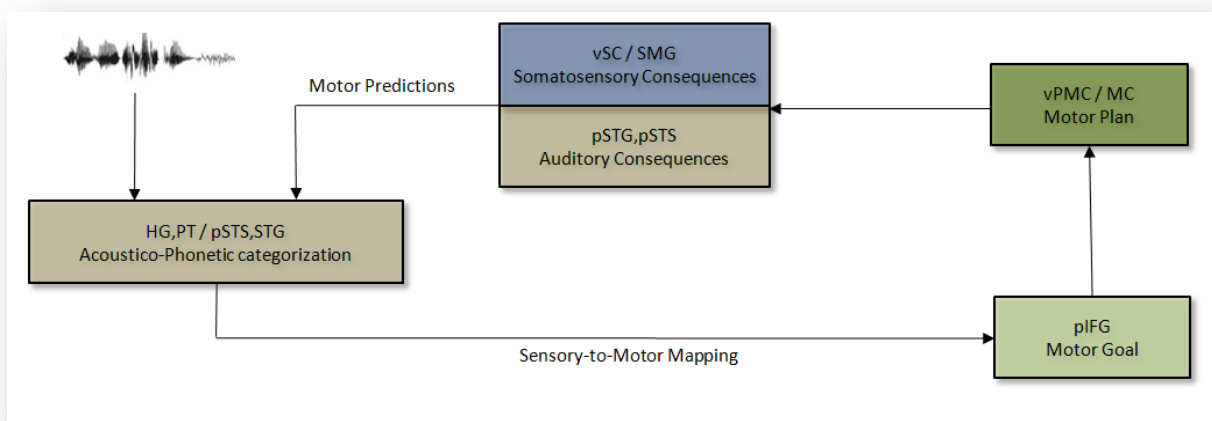


Figure I.12 : Schéma du modèle d'analyse-par-synthèse de Skipper et al. (2007). En cas d'ambiguïté perceptive, l'interprétation phonétique du signal de parole est considérée être effectuée par un appariement des conséquences sensorielles générées à partir de la simulation motrice de possibles candidats articulatoires.

Celles-ci vont être projetées sur le gyrus frontal inférieur et appariées aux buts articulatoires qui peuvent le plus typiquement être à l'origine de ces hypothétiques phonèmes. Ensuite, le cortex prémoteur ventral, en passant par le cortex moteur primaire, simule des commandes motrices sous-jacentes et de là, émet des copies d'efférence contenant les conséquences sensorielles prédites. Celles-ci seront renvoyées au cortex auditif afin de contraindre l'interprétation phonétique de l'hypothèse phonémique.

De manière capitale, les auteurs proposent une implication des centres moteurs de la parole en fonction de la modalité de présentation et de l'ambiguïté des entrées sensorielles, et ce d'autant plus que la correspondance entre information sensorielle et catégories phonétiques est dégradée.

La production de la parole sous l'angle des interactions sensorimotrices

Contrairement au rôle possible et débattu du système moteur lors de la perception de la parole, les interactions entre systèmes sensoriels et moteur sont largement considérées comme jouant un rôle fondamental en production de la parole. Ces interactions sont en effet supposées permettre une comparaison des conséquences sensorielles prédites des actes moteurs de parole et des entrées sensorielles réelles et ainsi un contrôle en ligne des productions de parole.

Dans ce sous-chapitre, un aperçu des études et résultats comportementaux ayant mis en évidence de telles interactions sensorimotrices lors de la production de parole sera suivi d'une description des bases cérébrales de la production et des modèles neurobiologiques récents postulant l'existence de mécanismes sensorimoteurs mis en œuvre dans la régulation et correction des actes de parole.

Apports des études comportementales

Etant donné le nombre quasi illimité de possibles comportements moteurs (compte tenu par exemple du nombre de configurations possibles de chaque articulateur par rapport à chaque degré de liberté; Mussa-Ivaldi, 1999), une représentation propre à chaque mouvement est difficilement concevable. Un concept central lié aux recherches portant sur le contrôle de l'action en général et sur le contrôle moteur lors de la production de parole en particulier est celui de "modèle interne". Inspirés par Francis et Wonham (1976), Kawato et collaborateurs (Kawato, Furokawa et Suzuki, 1987) ont été parmi les premiers à proposer un modèle interne d'actions motrices chez l'homme qui assure le bon déroulement de mouvements malgré des possibles perturbations en ajustant des paramètres de mouvement en fonction des besoins de l'action. Ils ont ainsi proposé un système de boucles régulatrices sensorimotrices lors du contrôle moteur des actions afin d'expliquer la capacité du système moteur humain à exécuter des actions motrices efficacement, malgré les délais des boucles de contrôle par feedback sensoriel trop long pour un contrôle en ligne d'un mouvement rapide et complexe (par exemple, le traitement du feedback auditif de la production de la parole serait de l'ordre de 200ms, Houde et al., 2002). Pour Wolpert, la notion de modèle interne est reliée à l'existence de représentations internes nécessaires au contrôle moteur des actions (Wolpert, Ghahramani et Jordan 1995; Wolpert, 1997) et basée sur des principes de codage de patterns similaires de comportements moteurs, encodés en mémoire procédurale suite à un apprentissage sensorimoteur par essais et erreurs (Ganesh et al., 2010).

Il a été proposé que les modèles internes soient constitués d'un modèle inverse qui permet de retrouver les commandes motrices apprises et relatives à la réalisation d'une action donnée et d'un modèle direct qui permet de prédire le résultat de ces commandes en estimant et comparant la position et la dynamique des articulateurs en mouvement avec les conséquences sensorielles des commandes motrices effectuées (voir Figure I.12). Ainsi, le système moteur peut, à l'aide du modèle

directe, inférer les conséquences sensorielles des actions réalisées à partir du modèle inverse sous forme de 'copie d'efférence'.

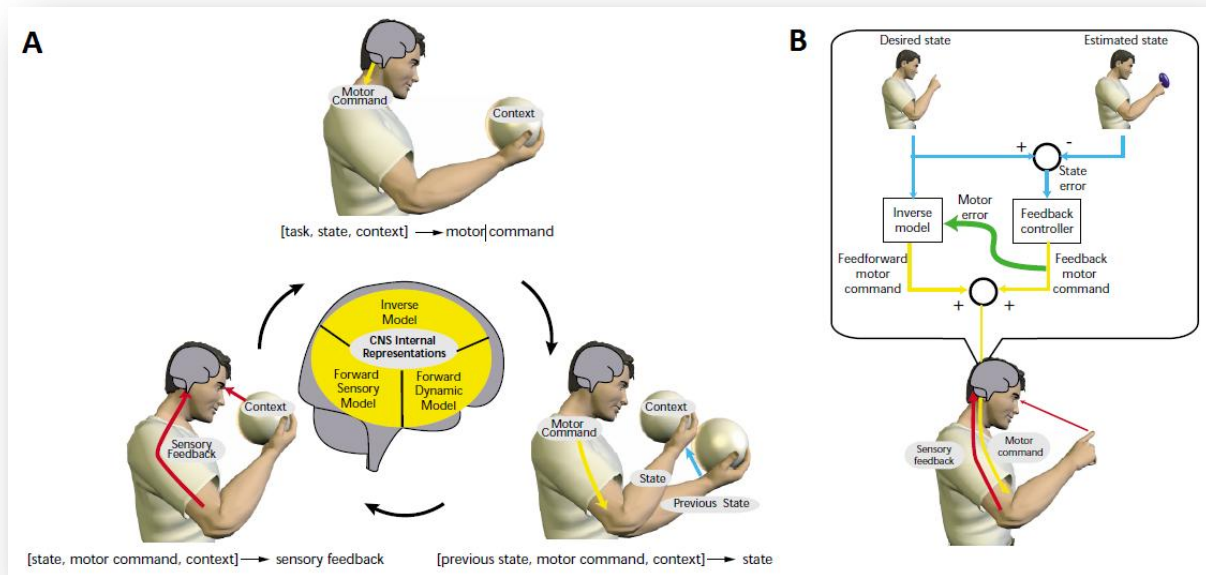


Figure 1.13 : (A) Boucle sensorimotrice lors de la réalisation d'une action montrant la génération des commandes motrices (haut) grâce au modèle inverse, un état de transition (droite) et la génération du feedback sensoriel (gauche). (B) Correction du mouvement (feedback motor command) et mise à jour du modèle inverse (motor error) lors de la réalisation d'une action par comparaison des conséquences sensorielles prédites et réelles liées à cette action (Figures empruntées à Wolpert et Ghaharamani, 2000).

Le concept de décharges corollaires²¹ (soit la transformation en prédictions sensorielles de la copie d'efférence) a été introduit à la suite d'expérimentations animales chez le poisson et l'insecte par Roger Sperry (1950) et Eric von Holst (1954). La décharge motrice produite par les centres nerveux et envoyée vers les muscles responsables des mouvements articulaires s'accompagnerait d'une décharge parallèle informant les systèmes perceptifs de la production de ces mouvements. Ce mécanisme permettrait ainsi une modification des mouvements en cas de conflit entre informations sensorielles prédites et réelles. D'autre part, il permettrait de distinguer une information due aux mouvements propres de l'individu, d'une information due à une modification du monde extérieur. En effet, lorsque la copie de la décharge motrice et l'information perceptive coïncident, cette dernière est reconnue comme auto-produite et par conséquent n'est pas prise en compte par le système nerveux comme une modification du monde extérieur. Un grand nombre de travaux ont depuis repris et étendu ce concept de copie d'efférence : par exemple dans le domaine de la vision portant sur l'invariance perceptive due à des mécanismes de suppression saccadique, celui de la planification motrice des mouvements dirigés de la main ou encore dans le domaine de la conscience et de l'agentivité par la description de modèles internes liés à la représentation de soi et d'autrui (pour des revues, voir par exemple, Wolpert, Ghaharamani et Jordan 1995; Blakemore, Wolpert et Frith 1998;

²¹ Les premières idées de copies d'efférence ont été proposées au 19^{ème} siècle en parallèle et indépendamment par Bell en 1823 et Purkinje en 1825, qui ont essayé de proposer une explication à la stabilité de l'image rétinienne : lorsqu'on appuie sur le globe oculaire les yeux fermés, l'image rémanente ne bouge pas, sauf en cas de mouvements des globes oculaires. Au contraire, l'image réelle bouge si on appuie sur le globe oculaire, mais pas si le globe oculaire est en mouvement pendant ce déplacement (pour plus d'explication et une discussion sur les copies d'efférence, voir Bridgeman, 2007). Par la suite, Hermann von Helmholtz a introduit l'idée de 'sensation de l'effort' pour expliquer le phénomène de stabilité perceptive par un mécanisme d'intégration sensorimotrice : lorsque les globes oculaires se déplacent, le système moteur envoie des signaux de 'sensation de l'effort' aux muscles oculaires pour compenser leur mouvement. De cette manière, malgré le mouvement des muscles oculaires, l'image sur la rétine resterait intelligible, car les paramètres de déplacement des globes oculaires auraient été intégrés pour l'interprétation du percept visuel (von Helmholtz, 1867/1910).

Jeannerod, 1994, 2001, 2003; Blakemore et Frith, 2003; Rapin, 2011; Perrier, 2012)²².

Appliqué à l'étude de la parole, le concept de modèles internes, direct et inverse, sous-tend la question de la nature des cibles et représentations invariantes motrices, auditives ou sensorimotrices des actes de parole. Cette question, complexe, est évidemment reliée à celle de la nature auditive, motrice ou sensorimotrice des objets perceptifs de parole et, plus largement, à la question des représentations des unités phonologiques.

A l'instar de la théorie motrice ou de la théorie directe réaliste de la perception de la parole (Liberman et al., 1962, 1967; Liberman et Mattingly, 1985; Fowler, 1986, 1996; Liberman et Whalen, 2000), de nombreux chercheurs postulent une nature motrice, articulatoire, du but des actes de parole. A titre d'exemple, citons Browman et Goldstein (1989) qui, dans le cadre de la théorie de la phonologie articulatoire, considèrent les gestes articulatoires comme primitives des représentations phonologiques et dont la coordination par couplage dynamique des cibles articulatoires représente le véritable but de l'action produite. Saltzman et collègues (1986) proposent également dans le modèle Task Dynamics que le but des actes de parole réside dans les trajectoires articulatoires, calculées à partir d'un paramètre spécifiant le point d'équilibre de la cible articulatoire, et intégrant la rapidité des mouvements et le degré de rigidité de la constriction.

A l'opposé, Joseph Perkell et ses collaborateurs au Speech Communication Group du MIT ont défendu la position de cibles auditives comme buts des actes de parole (pour une revue récente, voir Perkell, 2012; voir aussi Perrier, 2005). Puisque l'objectif premier du locuteur est de communiquer, et donc d'être intelligible, la production de patterns acoustiques reconnaissables par l'auditeur apparaît en effet primordiale. Nous rappellerons ici un certain nombre d'observations soulignant l'importance des représentations et buts auditifs lors de la production de la parole. Néanmoins, comme pour la perception de la parole, la question de la nature des cibles et représentations motrices, auditives ou sensorimotrices des actes de parole est complexe et reste largement débattue.

Effet d'équivalence motrice. Atal et collègues (1978) ont observé que la voyelle /u/ en anglais américain peut être réalisée articulatoirement de plusieurs manières : soit avec la langue plus élevée mais les lèvres moins arrondies, soit l'inverse. A partir de là, Perkell et al. (1992) ont proposé la notion de relation d'équivalence motrice ("motor-equivalent trading relation") qui tient en compte la relation entre la position de la langue et l'arrondissement des lèvres (pour réaliser la même voyelle /u/ il est possible d'utiliser les deux stratégies précédemment décrites et, également, tout un continuum de positions articulatoires tenant compte de cette relation langue-lèvres). Les travaux sur les perturbations motrices convergent également pour montrer que les locuteurs utilisent alors des stratégies de correction qui semblent, au moins partiellement, guidées par un objectif auditif (Savariaux, Perrier et Orliaguet, 1995; Savariaux et al., 1999).

Surdité et production de la parole. Des arguments empiriques forts en faveur de l'existence de buts acoustiques lors de la production de parole découlent de l'observation des productions de personnes sourdes de naissance. Il est connu que les personnes sourdes de naissance acquièrent difficilement une parole compréhensible et fluide (Cowie et Douglas-Cowie, 1983), alors que les personnes ayant acquis une surdité post-linguistique conservent au départ une production intelligible. Ces différences peuvent notamment s'expliquer du fait de l'apprentissage préalable à la surdité des relations entre buts moteurs, auditifs et somatosensoriels et l'utilisation des associations somatosensorielles et motrices une fois la surdité acquise (Matthies et al., 2008). Néanmoins, la finesse articulatoire des

²² Selon Wolpert et Kawato (1998) et Wolpert, Miall et Kawato (1998), le cervelet serait la structure apte à effectuer les transformations sensorimotrices prévues pour les modèles directs et inverses. En effet, le cervelet permet la coordination fine de mouvements grâce aux boucles sensorimotrices cortico-cérébelleuses. Les cellules Purkinje du cervelet reçoivent des copies d'efférence de la part des régions corticales motrices et sensorielles par les noyaux du pont (via les faisceaux fronto-pontique, pariéto-pontique ou temporo-pontique). Après articulation au niveau du noyau dentelé, les sorties peuvent être doubles : une partie des fibres sortantes font relai par les noyaux du thalamus pour retourner aux régions corticales et mettre à jour les commandes motrices. Une autre partie des fibres sortantes descendent dans le tronc cérébral afin de constituer les voies extrapyramidales d'ajustement moteur, les faisceaux réticulospinal et rubrospinal, toutes deux assurant la coordination des muscles striés.

productions de ces patients se dégrade avec le temps, notamment en termes de contrastes phonémiques (Cowie et Douglas-Cowie, 1992), un résultat parfois interprété en termes de détérioration du modèle direct (Perkell, 2012). Il a été également montré que l'utilisation d'implants cochléaires permet d'améliorer significativement la finesse de perception et de production de contrastes phonémiques et de variabilité articulatoire (Lane et al., 2007, Matthies et al., 2008), ce qui confirme l'importance du système auditif dans les mécanismes de contrôle de production en ligne de la parole (Perkell, 2012).

Perturbation du feedback auditif. Chez des sujets sains normo-entendants, il a été montré que la modification en ligne du feedback auditif modifie de manière quasi immédiate (quelques dizaines de millisecondes) la production. Ainsi, la présence d'un fort bruit ambiant, perturbant l'analyse du feedback auditif, provoque des modifications de production, par exemple une hyper-articulation et l'augmentation de l'intensité vocale (Lombard, 1911; Lane et Tranel, 1971; Houde et al., 2002). De même, l'induction d'un délai du feedback auditif perturbe les sujets dans leur production (Hashimoto et Sakai, 2003; Takaso et al., 2010). Enfin, la modification des valeurs acoustiques de la fréquence fondamentale (F_0 , correspondant à la hauteur tonale) et du premier formant (F_1 , correspondant à la première résonance du spectre acoustique) induit des mécanismes de compensation articulatoire. Ces effets d'adaptation/compensation vont dans le sens contraire de la modification du feedback acoustique et perdurent un certain temps après arrêt des modifications et retour à un feedback auditif normal (Jones et Munhall, 2000; Leung et Ciocca, 2011). Face à ces modifications articulatoires²³, il est intéressant de noter que ces mécanismes de compensation/adaptation motrice ne sont pas nécessairement volontaires et conscients mais plutôt involontaires et automatiques. Appuyant cette idée, une étude a en effet démontré le caractère irrépressible de ces compensations articulatoires même lorsque les sujets étaient prévenus des modifications du feedback acoustique et qu'il leur était demandé de ne pas modifier leur production (Munhall et al., 2009). Enfin, de manière importante, il a été montré que le degré de compensation articulatoire face à une modification du feedback acoustique était fonction de l'acuité auditive du sujet (Villacorta et al., 2007).

Adaptation auditive et convergence phonétique. Un dernier résultat appuyant l'existence de cibles auditives lors de la production de parole concerne les modifications articulatoires suivant la présentation répétée d'un stimulus auditif. Cooper et Lauritsen (1974) en exploitant le paradigme d'adaptation sélective de la parole ont en effet montré que l'écoute répétée de la syllabe occlusive non voisée /pi/ avait pour effet un raccourcissement du délai de voisement lors d'une tâche subséquente de production de la même syllabe²⁴. De manière similaire, des modifications de production de parole interviennent également dans le cadre d'interactions conversationnelles. Il est en effet établi l'existence de mécanismes de convergence et d'imitation apparaissant entre deux interlocuteurs lors d'une conversation. Ces mécanismes de convergence, soit la tendance d'un locuteur à imiter son interlocuteur, se manifestent au travers de la posture, des mouvements de la tête et des expressions faciales et, en ce qui concerne la parole, par le débit d'élocution, l'intensité vocale, la courbe mélodique ou encore une modification de la fréquence fondamentale et des formants vocaliques (pour une revue, Aubanel, 2011).

²³ De même, bien que l'apport des informations somatosensorielles soit moins étudié dans le cadre de la production de la parole, des travaux comportementaux impliquant des perturbations orofaciales mécaniques de la peau du visage et des muscles du conduit vocal ont aussi démontré des mécanismes de compensation articulatoire. Ces résultats soulignent donc également l'importance des buts somatosensoriels lors de la production de parole (par exemple, Gracco et Abbs, 1985; Tremblay, Shiller et Ostry, 2003; Nasir et Ostry, 2006).

²⁴ Eimas et Corbit (1973) ont pour la première fois utilisé le paradigme d'adaptation sélective pour étudier la détection des traits phonétiques, en présentant des stimuli issus d'un continuum acoustique lié à des paires de syllabes bilabiales (b,p), alvéolaires (d,t) ou vélaires (g,k), différenciées uniquement par le trait de voisement. Par rapport à une tâche d'identification sans adaptation préalable, ils ont démontré qu'une présentation répétée d'une syllabe impliquant une consonne à l'extrémité du continuum induisait un biais de reconnaissance vers la consonne de l'autre côté du continuum acoustique (voir aussi Chapitre 3.1).

Codage cérébral de la production de la parole

Trois systèmes de contrôle moteur de la parole

Du point de vue bio-mécanique et physique, la production de la parole implique trois systèmes différents travaillant de concert : le système respiratoire, le larynx avec les cordes vocales et le conduit vocal avec les articulateurs supralaryngés. De manière très simplifiée, tout acte moteur de parole est constitué de trois étapes. (1) La production d'un souffle pulmonaire par le système sous-glottique (diaphragme et muscles inspireurs et expirateurs) va expulser de l'air jusqu'au larynx. (2) La pression d'air va faire entrer en vibration les cordes vocales, qui avec le larynx sont à la source de la phonation et responsables de la production du flux laryngé (son périodique complexe). (3) Suite à la vibration des cordes vocales, l'air va remonter le long du pharynx, du conduit vocal et des cavités nasales qui vont moduler le flux laryngé et filtrer de cette manière l'air sortant par la bouche ou le nez.

La **respiration** est contrôlée par les centres respiratoires du tronc cérébral (qui transforment les signaux ascendants et descendant en patterns rythmiques), mais les centres corticaux de plus haut niveau rendent possibles la prise de conscience de cette activité, autrement automatique, pour la moduler si besoin et la déclencher volontairement en dehors du pattern respiratoire (par exemple dans le but d'effectuer des actions vocaliques, comme parler ou chanter) ou au contraire la suspendre (comme lors de la déglutition).

L'expiration est assurée par les muscles respiratoires abdominaux et thoraciques qui se contractent suite à la stimulation par les motoneurones phréniques et thoraciques de la corne ventrale dans la moelle épinière, alors que l'inspiration semble impliquer l'activation des motoneurones innervant le génioglosse. Deux centres générateurs de patterns (central pattern generator) dans le pont sont impliqués dans le maintien des rythmes respiratoires en coordonnant les phases respiratoires : les oscillations rythmiques des motoneurones crâniens impliqués dans la respiration sont gérées par le centre apneustique qui permet l'inspiration) alors que le centre pneumotaxique inhibe l'inspiration. De plus, deux groupes respiratoires bulbaires contrôlent le timing des mouvements d'inspiration ou la force de respiration. Un contrôle sensorimoteur du larynx (par les muscles laryngés innervés par le nerf laryngé supérieur) permet la protection des voies respiratoires en détectant la pression à l'aide des mécanorécepteurs. Il s'agit de contrôler la "poussée" inspiratoire et de gérer le timing de la respiration et diminuer la pression générée par les muscles thoraciques, comme le génioglosse, ainsi que de protéger les voies respiratoires supérieures par l'adduction des muscles laryngés (pour plus de détails, voir Sawczuk et Mosier, 2001; Jürgens 2002, 2009).

Pour le contrôle volontaire de la respiration, telle qu'utilisée dans les activités de phonation ou de parole, sont impliqués les centres corticaux et sous-corticaux (ganglions de la base et le cervelet) qui règlent les activités du tronc cérébral et impliquent de manière supplémentaire les centres de contrôle moteur du diaphragme (voir McKay et al., 2003).

Système laryngé et phonation. La phonation (ou vocalisation) nécessite la coordination du contrôle de la respiration et du système laryngé. Si la simple vocalisation est déclenchée par le cortex cingulaire chez les mammifères non humains et par l'aire motrice supplémentaire (chez l'humain), les commandes motrices de la phonation volontaire sont contrôlées par la voie pyramidale provenant du cortex moteur primaire, qui traverse le tronc cérébral, probablement dans la substance grise périaqueducale²⁵ (Jürgens, 2002), et sont ensuite transmises par la formation réticulée aux motoneurones phonatoires, qui déclenchent les mouvements musculaires du larynx. (Jürgens 2002, 2009).

²⁵ D'autres informations convergent dans cette structure, qui serait impliquée dans un ajustement plutôt que dans la génération de pattern de réponses : motivationnelles, mnésiques, émotionnelles, etc d'un grand ensemble de structures corticales. D'ici, les informations ne sont pas directement envoyées aux motoneurones, mais le "tri" est effectué par la substance réticulaire.

- Les muscles laryngés intrinsèques sont impliqués dans la phonation, ils permettent l'ouverture (muscles abducteurs, crico-aryténoïdiens postérieurs) et la fermeture (muscles adducteurs; l'aryténoïdien transverse qui rapproche les aryténoïdes l'un de l'autre et les crico-aryténoïdiens latéraux qui ferment la glotte lors de la phonation ce qui tend les cordes vocales) de la glotte ou bien de relâcher la tension des cordes vocales (muscles thyro-aryténoïdiens). Tous ces muscles sont innervés par le nerf récurrent. Enfin le muscle cricothyroïdien, qui étire les cordes vocales, est innervé par le nerf laryngé supérieur externe. Ces muscles sont contrôlés par les motoneurones du noyau ambigu dans le bulbe rachidien.
- Les muscles laryngés extrinsèques s'insèrent sur le larynx et le connectent avec les structures voisines (l'os hyoïde, le sternum et le pharynx). Pour la phonation, les trois muscles suivants jouent un rôle crucial : le cartilage thyroïde est connecté par le muscle thyrohyoïdien à l'os hyoïde et par le muscle sternothyroïdien au sternum, qui est connecté par le muscle sternohyoïdien à l'os hyoïde. De cette manière, l'ensemble de la structure est stabilisé de manière à ce que les muscles intrinsèques puissent s'exercer indépendamment. Les muscles extrinsèques permettent l'élévation et l'abaissement du larynx. Leur activité augmente lorsque la fréquence fondamentale du son émis s'éloigne de la valeur normale. Le muscle sternohyoïdien est sollicité lors de l'augmentation de la puissance sonore. Ces muscles sont innervés par l'anse cervicale, qui met en relation les motoneurones du noyau hypoglosse au niveau du bulbe rachidien et les motoneurones des premières cervicales, sur la partie ventrale de la moelle épinière.

Articulation supralaryngée. Les sons émis lors de la phonation sont modulés par la forme du conduit vocal, où le déplacement des articulateurs mobiles provoque des propriétés de résonance, qui déterminent les caractéristiques spectrales de chaque son de parole. Les articulateurs de la parole incluent les parties immobiles comme la cavité nasale, le nasopharynx, les dents, le palais dur, l'arête alvéolaire et des parties mobiles comme les cordes vocales, la langue, la mandibule, les lèvres, le velum (voile du palais), le pharynx ou le larynx. Les motoneurones contrôlant les différentes structures mobiles proviennent du noyau facial, trigéminal moteur, ambigu, hypoglosse ainsi que la corne ventrale des premières cervicales. Les muscles orofaciaux en lien avec la production de la parole impliquent entre autres :

- Pour des mouvements buccaux : les muscles orbiculaires et mentonniers permettent la protrusion des lèvres, alors que les muscles buccinateur, zygomatique et l'abaisseur de l'angle de la bouche sont nécessaires pour ramener les coins de la bouche de retour.
- Le degré d'abaissement de la mandibule est réglé par : les muscles digastriques (tout comme le génio-hyoïdien et mylo-hyoïdien) pour l'abaissement de la mandibule (lorsque l'os hyoïde est fixe) et le muscle masséter, principal élévateur et propulseur de la mandibule, avec les muscles pterygoïdiens (les muscles temporalis permettent plutôt la mastication).
- Les mouvements linguaux : intrinsèques (ils donnent la forme de la langue), composés d'un muscle longitudinal supérieur, et de trois paires de muscles : longitudinaux inférieurs, verticaux et transverses extrinsèques (ceux qui prennent naissance à l'extérieur de la langue mais se connectent sur elle, ils assurent la position de la langue) incluant le génioglosse (qui permet la protrusion et la dépression de la partie centrale de la langue), l'hyoglosse (qui permet la dépression de la langue), le styloglosse (qui permet l'élévation et la rétraction de la langue) et le palatoglosse (qui permet la dépression du palais mou, le déplacement des piliers antérieurs et l'élévation de la partie postérieure de la langue)²⁶.

²⁶ Par exemple, la position de la langue pour produire la voyelle /i/ nécessite que le muscle génioglosse effectue une légère protrusion de la langue en avant et en haut. Pour produire un /u/, la langue doit être en position relevée en arrière, grâce au muscle styloglosse, et pour produire un /a/, le hyoglosse doit ramener la langue en bas (exemples tirés de Jürgens, 2002).

- Enfin, l'abaissement du vélum, qui en position abaissé permet le passage d'air dans le nez pour obtenir des sons de parole nasalisés, s'effectue notamment par le muscle élévateur du voile du palais.

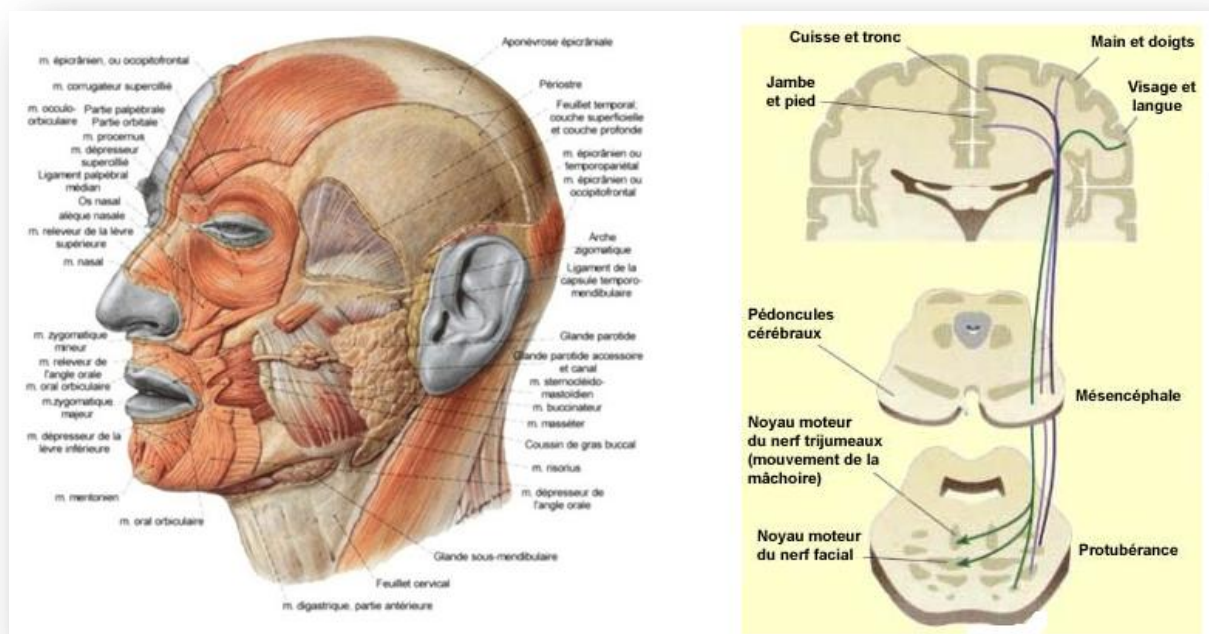


Figure I.14 : Gauche : Musculature orofaciale. Droite : voie pyramidale de contrôle moteur orofacial (source des figures : lecerveau.mcgill.ca).

Bases cérébrales du contrôle moteur de la parole

Contrôle moteur de la production de parole. Afin de produire des patterns vocaux appris, tels que les sons de parole dont l'objectif est d'effectuer des mouvements du conduit vocal très fins à des visées communicatives, il est nécessaire de contrôler à la fois la respiration, la phonation et la musculature supralaryngée. Liée initialement à un apprentissage sensorimoteur des représentations phonologiques, cette action complexe nécessite le fonctionnement de plusieurs boucles de contrôle impliquant des structures corticales et sous-corticales (pour des revues, voir Jürgens, 2002, 2009; Riecker, 2005).

- Une boucle de **préparation**, en charge des processus d'initiation et de coordination motrice (Riecker et al., 2005), serait composée du gyrus frontal inférieur et du cortex prémoteur (qui se chargent de l'élaboration et le séquençage du plan moteur), de l'insula²⁷, de l'aire motrice supplémentaire²⁸ et du cervelet supérieur. Il est intéressant de noter que des recherches plus récentes ont montré qu'il y aurait un processus de l'intention d'action, de choix de prise de parole (Carota et al., 2010) qui serait antérieur à et indépendant de l'élaboration du plan moteur de l'acte de parole²⁹.

²⁷ Les travaux de Dronkers (1996) ont montré une relation des lésions insulaires avec les troubles apraxiques de la parole, ce qui suggère la participation de cette structure dans la planification motrice des articulateurs orofaciaux pour la production de la parole.

²⁸ Pour les patterns vocaux appris chez l'humain, le gyrus cingulaire serait également impliqué dans l'initiation des actes de vocalisation, surtout en ce qui concerne les aspects émotionnels ou des patterns de vocalisation innés (voir Jürgens, 2002).

²⁹ Les mêmes types de conclusion ont été atteints par d'autres travaux ne portant pas sur la parole mais sur les gestes moteurs en général (Desmurget et al., 2009; Desmurget et Sirigu, 2009). Les auteurs ont stimulé directement le cortex pariétal inférieur et ont observé des envies ou des illusions de bouger différentes parties du corps, sans pour autant les effectuer réellement. Indépendamment de cela, les stimulations des régions prémotrices avaient pour conséquences des exécutions de mouvements, sans prise de conscience de ces actes par les patients.

- Une boucle **d'exécution** motrice inclurait le cortex moteur primaire orofacial (la première étape de la voie pyramidale), le thalamus (le principal relai entre le cortex et les autres structures centrales), le putamen et le noyau caudé, ainsi que le cervelet inférieur. Les fibres provenant des aires motrices primaires des deux hémisphères arrivent via la capsule interne et la couronne rayonnante, traversent le thalamus et, arrivées au niveau de la face ventrale du tronc cérébral ventral, décussent dans leur majeure partie au niveau des pyramides bulbaires. Les fibres (provenant des deux hémisphères) traversent la formation réticulaire pour atteindre les noyaux moteurs et ensuite principalement la voie corticobulbaire qu'empruntent les nerfs crâniens (V trijumeau, VII facial, IX glosso-pharyngien, X vague, XI accessoire et XII hypoglosse) pour le contrôle de la musculature supralaryngée, phonatoire et respiratoire. Une partie de la musculature laryngée et respiratoire (thoraciques et abdominaux) requiert le contrôle de mouvements volontaires par les nerfs spinaux (au niveau des segments thoraciques et abdominaux) via la voie corticospinale. Il est à noter qu'une fois sollicité par l'aire motrice supplémentaire, le cortex moteur primaire doit obligatoirement faire relai par la formation réticulée afin d'atteindre les motoneurones de vocalisation et c'est au niveau de la formation réticulée qu'aurait lieu l'intégration des composantes respiratoire, laryngée et articulaire (Jürgens, 2009). Les ganglions de la base forment en soi cinq boucles sensorimotrices interconnectées, qui permettent l'exécution du but moteur en régulant les mouvements agonistes et antagonistes via l'inhibition de mouvements involontaires et la facilitation de mouvements prévus (voir Obeso et al., 2000 pour plus de détails). Le cortex moteur de la face projette sur le putamen ventral, qui transmet alors l'information à la substance noire et, de là, à la formation réticulée parvocellulaire, directement connectée aux motoneurones phonatoires (Iwata et al., 1996)³⁰.
- Un troisième ensemble fonctionnel de **correction/régulation motrice** peut être proposé par rapport aux travaux sur le contrôle en ligne de la production de la parole (discutés plus en détail ci-dessous) et inclurait le cortex somatosensoriel, le lobule pariétal inférieur et, dans certains cas, le gyrus temporal supérieur (voir aussi Dhanjal et al., 2008 et Peschke et al., 2009). Les afférences sensorielles (conséquences somatosensorielles et auditives des commandes motrices exécutées) arrivent via le thalamus ventrolatéral et convergent dans le cervelet au niveau des cellules de Purkinje, où elles sont mises en correspondance avec les copies d'efférence des prédictions sensorielles en conséquences des commandes motrices envoyées par les régions sensorimotrices corticales, ce qui permet l'élaboration d'éventuelles corrections. Celles-ci vont quitter le cervelet par les noyaux interposé et dentelé via les pédoncules cérébelleux pour traverser le thalamus ventrolatéral et transmettre une possible mise à jour de la commande motrice aux régions motrices (Nakano et al., 1992). Selon différents modèles, ces corrections seront amenées aux structures exécutrices (aire motrice primaire, aire motrice supplémentaire) ou planificatrices (aire de Broca, cortex prémoteur).

Interactions sensorimotrices en production de la parole

Comme décrit précédemment, l'existence d'interactions sensorimotrices est un concept central dans le domaine de la production de la parole. Du point de vue cérébral, une modulation des réponses du cortex auditif lors de la production ouverte de parole, par rapport aux réponses neuronales observées lors de l'écoute passive de ces mêmes productions, a été démontrée dans de nombreuses études utilisant les techniques d'électroencéphalographie, de magnétoencéphalographie et d'IRMf (voir par exemple Paus et al., 1996; Numminen et Curio, 1999; Curio et al., 2000; Houde et al., 2002; Christoffels, Formisano et Schiller, 2007; Christoffels et al., 2011). Pris ensemble, ces résultats sont généralement interprétés dans le cadre théorique de modèles internes direct et inverse (voir ci-dessus). La modulation de réponse du cortex auditif est en effet supposée refléter des mécanismes de contrôle du feedback acoustique permettant d'évaluer les conséquences auditives prédites des

³⁰ Pour une revue du rôle des ganglions de la base dans le contrôle moteur, voir Brooks, 1995. Pour une lecture concise des troubles neurologiques du contrôle moteur de la parole, voir Pinto et Ghio, 2007.

retours réels des actes moteurs de manière à contrôler et corriger si nécessaire notre production et de distinguer les conséquences sensorielles de nos actions des signaux sensoriels de notre environnement (pour des revues, voir Guenther, 2006; Guenther et Vladusich, 2012).

De même, une modulation de l'activité du cortex auditif a été observée lors de la modification en ligne du feedback auditif. Ce résultat est observé par l'ajout de bruit (par exemple, Christoffels, Formisano et Schiller, 2007; Christoffels et al., 2011; Zheng, 2010), en effectuant un décalage temporel du feedback auditif lors de la prononciation de mots ou de phrases (Hashimoto et Sakai, 2003; Takaso et al., 2010) ou en modifiant directement les propriétés acoustiques du signal de parole produit. Par exemple, par rapport à une condition sans perturbation auditive, une diminution d'amplitude du potentiel évoqué auditif N1 (EEG; Heinks-Maldonado et al., 2005) et M100 (MEG; Heinks-Maldonado et al., 2006), apparaissant 100ms après la production de la voyelle, a été observée dans les régions auditives bilatérales suite à une modification en ligne de la fréquence fondamentale (F_0) de voyelles produites (pour des résultats similaires lors de la production de mots, voir Fu et al., 2005). En accord avec une spécialisation hémisphérique de l'hémisphère droit dans le traitement des informations prosodiques et la distinction des hauteurs tonales, Toyomura et al. (2007) ont modifié la fréquence fondamentale lors de la production soutenue de la voyelle [a] et ont observé des activations accrues dans le gyrus temporal supérieur, le sulcus pariétal inférieur et le gyrus supramarginal, le cortex préfrontal, l'insula antérieure de l'hémisphère droit (ainsi que dans le cortex prémoteur de l'hémisphère gauche). Tourville, Reilly et Guenther, (2008) ont mesuré les modulations neuronales lors de la modification en ligne du premier formant (F_1) lors de la production de syllabes et ont observé une activation accrue au niveau des gyri temporaux supérieurs bilatéraux (tout comme l'activation des aires prémotrices et motrices ventrales et le cervelet antérieur, à droite). Enfin, de manière intéressante, Golfinopoulos et al. (2011) ont montré un effet similaire de perturbation de la production de la parole en modifiant le feedback somatosensoriel (en bloquant de manière imprévue les mouvements de la mâchoire). Par rapport à une condition contrôle sans modification somatosensorielle, ils ont observé des activations accrues du cortex moteur primaire et prémoteur ventral, du gyrus frontal antérieur, du gyrus supramarginal et du cervelet, ainsi qu'un renforcement des connexions fonctionnelles entre ces régions.

Modèles neurobiologiques de production de la parole

Pris ensemble, ces résultats sont généralement interprétés dans le cadre de mécanismes de contrôle articulatoire et de boucles sensorimotrices régulatrices. Par exemple, dans le modèle (neuro)computationnel DIVA de production de la parole (Directions Into Velocities of Articulators; Guenther, 2006; Guenther et Vladusich, 2012), la modulation des réponses neuronales du cortex auditif est considérée refléter des mécanismes de contrôle et de comparaison en ligne des représentations articulatoires et auditives. Les conséquences sensorielles prédites de l'acte moteur seraient ainsi évaluées par rapport au feedback sensoriel réel de nos actions de manière à permettre, si nécessaire, des modifications rapides de production de parole.

Modèle DIVA de la production de la parole. Le modèle DIVA de production de la parole développé par Guenther et collaborateurs (par exemple, Guenther, 2006; Guenther et Vladusich, 2012; voir Figure 1.15) est une implémentation neuro-computationnelle principalement dérivée des propositions théoriques de Perkell et collaborateurs citées précédemment. L'intérêt de ce modèle est, d'une part, de tenir compte de plusieurs propriétés caractéristiques de la production de parole (comme les principes d'économie articulatoire, de co-articulation, d'équivalence motrice ou encore les étapes d'acquisition de l'inventaire phonémique et la prise en compte de mécanismes de régulations sensorimotrices) et, d'autre part, de s'appuyer sur de nombreuses vérifications expérimentales comportementales et neurophysiologiques.

Le modèle DIVA comprend deux systèmes de contrôle direct (feedforward) et par rétroaction (feedback), qui se co-structurent lors de l'apprentissage du modèle (simulation des étapes d'acquisition de la parole lors du développement). Afin de pouvoir efficacement gérer les

commandes motrices en fonction des buts sensoriels du locuteur, le système direct doit apprendre lors de cette étape les relations entre représentations phonémiques, motrices, auditives et somatosensorielles. Les cibles de parole sont initialement de nature auditive et guidant, grâce au système rétroactif, la mise en place des commandes motrices sous-jacentes. Suite à cette phase d'apprentissage, des cibles somatosensorielles sont créées et intégrés dans le système de contrôle.

Chez l'adulte et pour des unités surappries, ces deux systèmes peuvent fonctionner de manière relativement indépendante et parallèle. Le système direct serait impliqué dans l'exécution des commandes motrices tandis que le système rétroactif, dont le rôle est de comparer les prédictions des actes moteurs avec les conséquences sensorielles réelles, n'interviendrait qu'en cas de situations conflictuelles (par exemple, dans des conditions environnementales bruitées).

Dans ce modèle, la production d'une unité de parole commence par l'activation de la carte des sons de parole ("speech sound map"), supposée contenue dans la partie operculaire du gyrus frontal inférieur gauche et le cortex prémoteur ventral adjacent. Cette carte des sons de parole répertorie tous les sons appris par le modèle dans la phase d'apprentissage (un phonème, une syllabe ou même une séquence de syllabe, voir Guenther et Vladusich, 2012). Bien que reliées à des structures prémotrices, l'emploi des termes de cartes de sons de parole pour désigner cette unité de stockage reflète la nature supposée sensorielle, auditive et somatosensorielle, des cibles de parole.

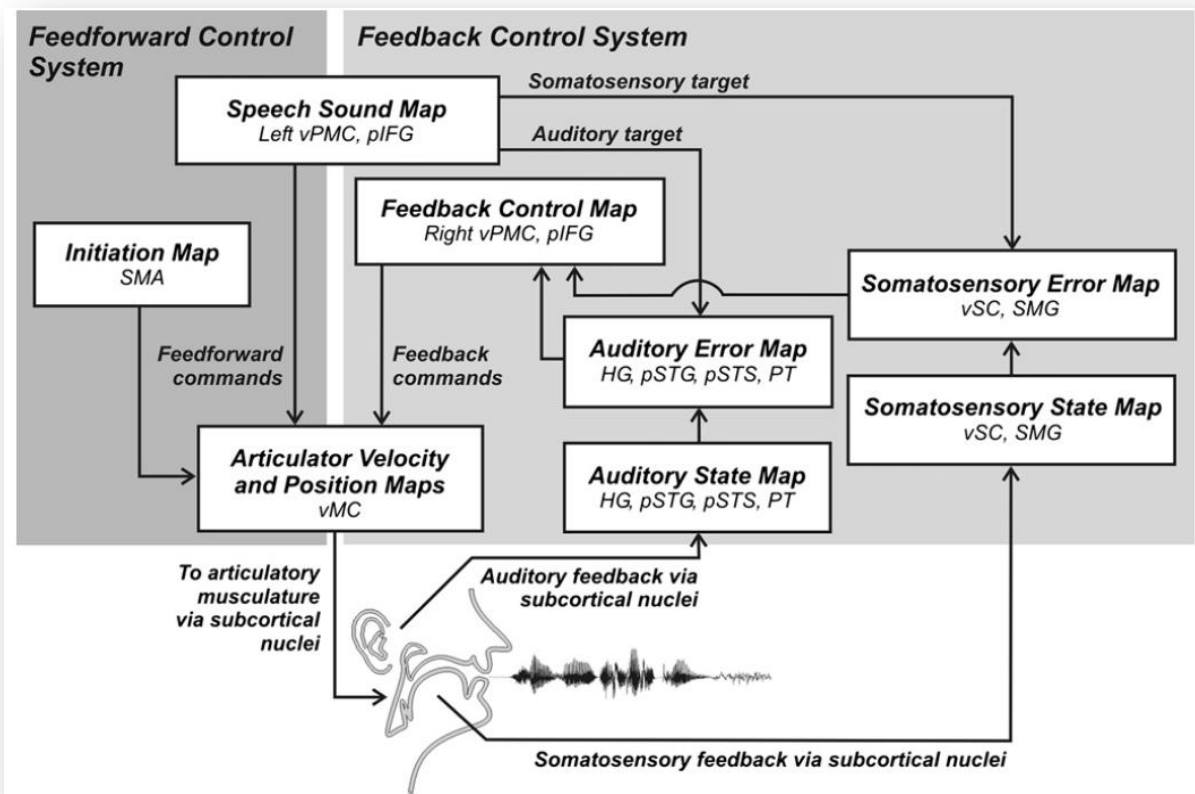


Figure I.15 : Illustration schématique du modèle DIVA de production de la parole avec les composantes cérébrales correspondant aux systèmes direct et rétroactif. Les représentations phonémiques siègent au niveau du cortex prémoteur et du gyrus frontal inférieur gauches. Lors de la production d'une unité de parole, les commandes motrices sont envoyées au système direct d'exécution situé au niveau du cortex moteur primaire. En parallèle, les conséquences sensorielles des commandes motrices sont envoyées au système de contrôle rétroactif (cartes d'erreurs auditives et somatosensorielles). Les cartes d'erreurs comparent ces prédictions sensorielles avec les signaux afférents et, en cas d'erreurs, transmettent, via le gyrus frontal inférieur postérieur et le cortex prémoteur de l'hémisphère droit, les informations nécessaires au cortex moteur primaire pour une modification adéquate des commandes motrices (Figure empruntée à Guenther et Vladusich, 2012).

A partir de là, les deux systèmes sont mis en route. En ce qui concerne le système direct, le cortex moteur primaire reçoit les informations de l'unité à produire sous forme de plan moteur (élaboré par le cervelet et stocké dans la carte des sons de parole) de manière à exécuter le programme moteur requis (via la voie pyramidale). En parallèle, le système rétroactif reçoit une copie de la commande motrice, ou plus exactement, les conséquences/cibles sensorielles de celle-ci, au sein de cartes d'erreurs auditives (situées dans le cortex temporal supérieur postérieur bilatéral) et somatosensorielles (situées dans les aires somatosensorielles primaires orofaciales et le gyrus supramarginal bilatéral). Les conséquences sensorielles réelles sont analysées par des cartes d'état sensoriel auditif et somatosensoriel (dans le cortex temporal supérieur postérieur bilatéral et les aires somatosensorielles primaires orofaciales et le gyrus supramarginal bilatéral). En cas d'une détection d'erreurs entre conséquences réelles et prédites³¹, les informations nécessaires à la correction de l'acte moteur sont envoyées à une carte de contrôle rétroactif (située dans la partie operculaire du gyrus frontal inférieur droit et le cortex prémoteur ventral adjacent) qui, enfin, signale au cortex moteur primaire les corrections nécessaires.

Autres modèles de production de la parole. Des modèles similaires faisant référence à des modèles internes du conduit vocal ont été également proposés.

Modèle de prédiction en série de Tian et Poeppel (2010). Ce modèle comporte 4 sous-systèmes principaux. (1) Le cortex pariétal émet l'intention d'un mouvement à effectuer et son plan moteur est élaboré par le système moteur. (2) A partir de là, deux informations sont envoyées en parallèle : la commande motrice est envoyée à la périphérie pour la réalisation de l'acte de parole et une copie d'efférence de cette commande motrice est envoyée (3) à deux modèles directs. Le premier modèle direct estime l'état moteur et les paramètres articulatoires. Une seconde copie d'efférence (à partir des estimations du premier modèle direct moteur) est envoyée 170ms plus tard à un deuxième modèle direct qui prédit les conséquences sensorielles de ces estimations articulatoires. A partir de là, (4) l'état moteur est mis à jour en ajustant les paramètres articulatoires aux estimations motrices et en confrontant les prédictions avec les retours sensoriels réels³².

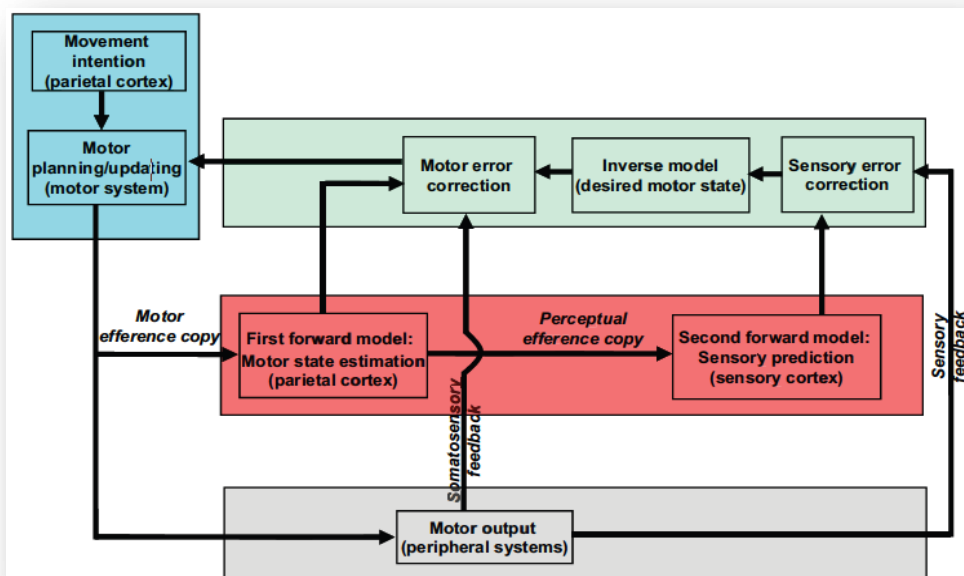


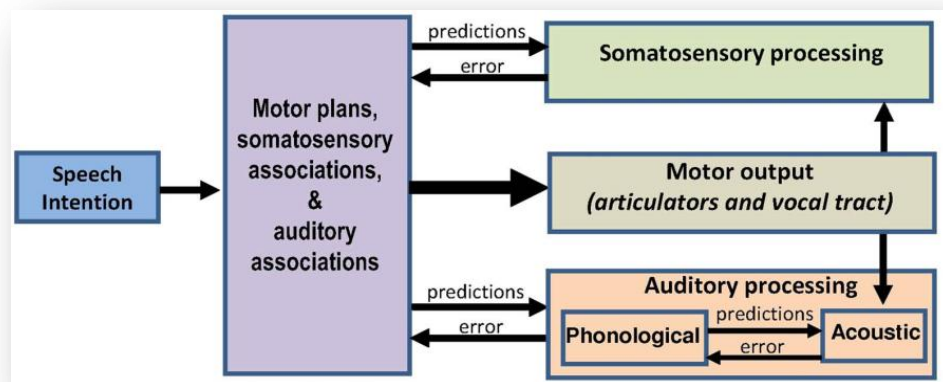
Figure 1.16 :
Modèle de
prédiction en
série de Tian
et Poeppel
(2010).

³¹ L'activation accrue des régions auditives et somatosensorielles lors de la production témoignerait d'une signalisation par les cartes d'erreurs d'incohérences entre informations sensorielles prédites et réelles, alors que l'absence de modulation d'activité de ces régions indiquerait une absence d'erreurs et/ou la non-implication de ce système de contrôle rétroactif, par exemple lors de la production d'unités de parole surappriées dans des conditions d'écoute favorables.

³² A noter que dans le modèle DIVA des traitements prédictifs sensoriels similaires sont effectués en parallèle aux traitements des commandes motrices, mais non pas en série comme dans le présent modèle de Tian et Poeppel.

Modèle génératif de Price, Crinion et MacSweeney (2011). L'aire de Broca et le sulcus temporal supérieur postérieur seraient impliqués dans la prédiction des conséquences sensorielles (auditives) de l'articulation, et ce, même en l'absence de feedback auditif. Ce modèle s'inspire des principes génératifs prédictifs, qui consistent en une structure hiérarchique dans laquelle les systèmes de plus haut niveau prédisent les entrées des systèmes de plus bas niveaux, ces prédictions pouvant ensuite être exploitées pour l'optimisation de prédictions futures. Ainsi, grossièrement, les sites de plus haut niveau prédisent les associations phonologiques des mots qu'elles codent, et par la suite, le système de traitement phonologique prédit les conséquences auditives correspondantes. Les éventuelles erreurs de prédiction sont renvoyées aux sources des prédictions (les sites de traitement de haut niveau), celles-ci étant en charge de remodeler les représentations haut niveau du mouvement, et, à partir de celles-ci, d'émettre de nouvelles prédictions. Ainsi, l'aire de Broca code les représentations articulatoires de haut niveau des mots surappris, ce qui permet de prédire les conséquences auditives de l'articulation de ces mots. Ces prédictions seraient envoyées aux régions temporales incluant le planum temporale et le sulcus temporal supérieur postérieur (sites d'associations multimodales). Ainsi, elles permettraient d'effectuer des prédictions auditives (sur la base des prédictions phonologiques) en termes d'associations acoustiques des mots articulés. Plus particulièrement, c'est le sulcus temporal supérieur postérieur qui élaborerait les représentations auditives phonologiques correspondant aux mots articulés, alors que la fonction du planum temporale ne serait pas spécifique à la parole mais jouerait plutôt un rôle d'association acoustique à n'importe quel type de mouvement.

Figure 1.17 :
Modèle génératif
de contrôle
moteur de Price,
Crinion et
MacSweeney
(2011).



Modèle hiérarchique de contrôle moteur par feedback de Hickok (2012). Ce modèle propose d'intégrer de manière hiérarchique les aspects psycholinguistiques et moteurs de la production de la parole. Le modèle reprend l'idée d'un système conceptuel et lexical (pour une revue, voir Indefrey et Levelt, 2004) et, de là, propose une décomposition des mots produits en phonèmes et syllabes. Ce modèle se base également sur la supposition que les buts des actes de parole sont de nature sensorielle et codés à différents niveaux hiérarchiques : les buts de "haut niveau" correspondraient aux structures syllabiques et seraient codés par des représentations auditives, alors que les buts de "bas niveau" correspondraient aux phonèmes et seraient codés en termes somatosensoriels. Lors de la réalisation des phonèmes et syllabes, des projections seraient envoyées en parallèle au système moteur et aux systèmes sensoriels, pour dans ce cas permettre une comparaison des prédictions auditives et somatosensorielles avec les retours réels. Pour les syllabes, ces transformations audiomotrices entre programme moteur syllabique (codé dans la partie postérieure du gyrus frontal inférieur) et représentations auditives correspondantes (partie postérieure du gyrus supérieur temporal) seraient effectuées grâce à une région à la frontière temporo-pariétale au niveau de la scissure sylvienne (l'aire 'SPT'). Pour les phonèmes, les transformations somatosensorimotrices entre plan moteur phonémique (cortex prémoteur ventral et aire motrice primaire) et représentations auditives (cortex somatosensoriel primaire et partie antérieure du gyrus supramarginal) seraient

réalisées par le cervelet. Les éventuelles erreurs de prédictions seraient corrigées de manière à mettre à jour les commandes motrices relatives aux syllabes et aux phonèmes produits.

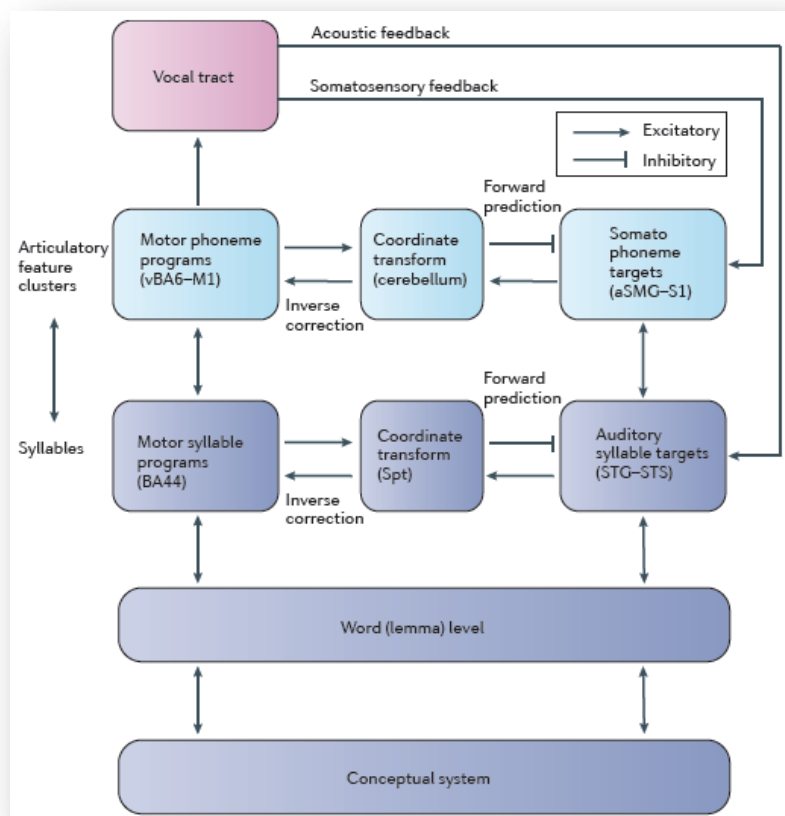


Figure I.18 : Modèle hiérarchique de contrôle moteur par feedback de Hickok (2012).

Problématique de la thèse et guide de lecture

L'hypothèse principale défendue dans la présente thèse est celle d'un couplage fonctionnel des systèmes de perception et de production de la parole au travers de la coordination des représentations auditives, somatosensorielles et motrices sous-jacentes. Dans ce cadre, nous avons tenté de mieux cerner la nature des représentations cérébrales sous-tendant la parole par l'examen des activités sensorielles et motrices, communes et distinctes, lors de la perception et de la production d'actes de parole, et d'un possible codage distribué sensorimoteur de ces activités.

La première partie de ce travail avait pour objectif, par l'utilisation de l'imagerie par résonance magnétique fonctionnelle, une exploration fine des réseaux neuronaux et des cartes sensorielles et motrices activés aussi bien lors de la réalisation de gestes articulatoires orofaciaux supralaryngés que lors de la production et de la perception des voyelles du Français. Ces études devaient permettre de clarifier les réseaux neuronaux sous-jacents au contrôle moteur minimal orofacial liés à des mouvements simples supralaryngés et laryngés/vocaliques, d'affiner l'organisation somatotopique des représentations motrices et somatosensorielles labiales, mandibulaires, linguales et vocaliques/laryngées et linguales et de déterminer un possible codage distribué au sein des régions sensorielles et motrices de différentes voyelles du Français, perçues et produites. Enfin, l'utilisation des voyelles devait permettre de déterminer un possible couplage fonctionnel des systèmes de perception et de production de la parole et ce, pour des unités élémentaires de la parole, spécifiées et décrites en termes articulatoires et acoustiques précis autour d'un jeu universel de traits phonétiques et peu contaminées par des effets complexes de co-articulation.

Face aux résultats observés, nous avons utilisé dans la seconde partie de nos travaux un paradigme d'adaptation lors de la mesure de l'activité neuronale par imagerie par résonance magnétique fonctionnelle afin d'identifier les régions neurales impliquées dans le codage des actions produites ou perçues et reliées à des mécanismes d'adaptation sensorimotrice et ce, lors de la production et de la perception répétée des mêmes mouvements orofaciaux supralaryngés et des voyelles du Français.

Les études présentées dans la troisième partie de cette thèse avaient pour but de déterminer l'implication fonctionnelle et causale des régions sensorielles et motrices lors de la perception d'unités de parole, par l'utilisation de la technique de stimulation magnétique transcrânienne et de paradigmes d'apprentissage moteur et auditif à court-terme.

Enfin, ces résultats observés sur les corrélats neuronaux des gestes articulatoires et des voyelles, sur les buts sensorimoteurs des actions orofaciales supralaryngés et des voyelles perçues et produites ainsi que sur le rôle causal du système moteur lors de la perception des unités de la parole seront discutés et mis en perspective dans la partie finale de cette thèse.

Il est à noter que les parties expérimentales présentées dans cette thèse sont organisées selon une logique linéaire et progressive du point de vue de la "complexité" des systèmes étudiés (orofaciaux et vocaliques) ainsi que des niveaux de description (réseaux neuronaux et codage des représentations, buts des actions perçues et produites et rôle fonctionnel des régions sensorimotrices). De manière à permettre une lecture relativement indépendante de ces différentes parties expérimentales, les enjeux et bases théoriques de ces travaux seront rappelés dans les parties introductives. De fait, nous prions le lecteur de nous excuser pour certaines redondances inévitables.

- 1^{ère} Partie -

Les cartes sensorimotrices
de la parole

Avant-propos

Les deux études présentées dans ce premier chapitre avaient pour but, par l'utilisation de l'imagerie par résonance magnétique fonctionnelle, une description des structures cérébrales liées au contrôle moteur orofacial de mouvements simples supralaryngés linguaux, labiaux, mandibulaires et laryngés vocaliques (Chapitre 1.1) ainsi que des régions neurales associées à la perception et à la production des voyelles du Français (Chapitre 1.2).

Outre une description neuroanatomique et fonctionnelle de ces réseaux, l'objectif principal de nos travaux était de tester un possible couplage fonctionnel des systèmes de perception et d'action. Face aux travaux portant sur des unités de parole plus complexes, l'utilisation des voyelles du Français, considérées comme unités élémentaires de la parole, devait permettre de préciser dans ce cadre minimal la possible nature sensorimotrice des représentations cérébrales de la parole.

Concernant l'implémentation neuronale de ces unités de parole, nous avons tenté de préciser la cartographie cérébrale des gestes orofaciaux et des unités de parole perçues et produites, par l'examen d'une somatotopie des articulateurs de la parole au sein des aires motrices et somatosensorielles primaires et d'une possible topographie spécifique, une phonémotopie, des voyelles ou des traits phonétiques associés au sein des régions auditives, motrices et somatosensorielles.

Ces deux études ont été réalisées en collaboration avec Marc Sato, Jean-Luc Schwartz, Coriandre Vilain, Nathalie Vallée (Département Parole & Cognition, Laboratoire GIPSA-lab, UMR CNRS 5216 & Université de Grenoble), Laurent Lamalle, Irène Troprès, Jean-François Le Bas (Centre Hospitalier Universitaire de Grenoble, Structure Fédérative de Recherche N°1 "RMN Biomédicale et Neurosciences", Unité IRM 3T) et Monica Baciú (Laboratoire de Psychologie et Neurocognition, UMR CNRS 5105 & Université Pierre Mendès France).

Ces études ont été principalement subventionnées par un financement de l'Institut National Polytechnique de Grenoble (Bonus Qualité Recherche "Modyc", "Modélisation dynamique de l'activité cérébrale") et par des financements de l'Agence Nationale de la Recherche (ANR "SPIM", "L'imitation dans la parole : des intégrations sensorimotrices à la dynamique des interactions conversationnelles") et de l'Association Française de la Communication Parlée (AFCP, participation aux XXVIIIèmes Journées d'Étude sur la Parole, Mons, Belgique).

Publications pertinentes

- Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M. (2012). Functional MRI assessment of orofacial articulators: neural correlates of lip, jaw, larynx and tongue movements. *Human Brain Mapping*.
- Grabski K, Schwartz JL, Lamalle L, Vilain C, Vallée N, Baciú M, Le Bas JF & Sato M. (soumis). Shared and distinct neural correlates of vowel perception and production. *Journal of Speech, Language & Hearing Research*.

Considérations méthodologiques

Ce chapitre a pour but de décrire le protocole expérimental général ainsi que les méthodes d'acquisition et de prétraitement des images fonctionnelles et anatomiques communes aux deux études IRMf présentées dans la première partie de cette thèse. L'enregistrement des données relatif à ces études a en effet eu lieu lors d'une même session expérimentale et concernait le même groupe de participants. Il est également à noter que les études IRMf présentées dans la seconde partie de cette thèse et réalisées auprès d'un second groupe de participants impliquaient des protocoles expérimentaux et des méthodes d'acquisition et de prétraitement des données fonctionnelles et anatomiques en tous points identiques, à l'exception de l'utilisation d'un paradigme d'adaptation. Ces méthodes d'acquisition et de prétraitement ne seront plus par la suite présentées lors de la présentation des études IRMf décrites dans les deux premières parties de cette thèse.

Protocole expérimental

Recrutement

Les participants ont été recrutés selon les critères d'inclusion suivants : être de langue maternelle française, droitiers (selon le test de latéralité de Oldfield, 1971), ne pas avoir d'antécédents de troubles moteurs, auditifs ou langagiers et avoir une vue normale ou corrigée. Les participants étaient de plus examinés par un médecin neurologue en regard d'éventuels troubles neurologiques, psychiatriques ou autres troubles médicaux et contre-indications à l'IRM (voir Annexe 1 pour une description complète des critères d'inclusion et d'exclusion). Suite à cet examen, tous les participants ont donné leur consentement éclairé.

Ces études ont reçu un avis favorable du Centre Hospitalier Universitaire de Grenoble, du Comité de Protection des Personnes pour la Recherche Biomédicale de Grenoble et de l'Agence Française de Sécurité Sanitaire des Produits de Santé (N° CHU Promoteur : 0824; N° AFSSAPS : 2008- A01618-47). Ces études ont été effectuées selon les standards éthiques de la déclaration d'Helsinki de 1964. En compensation de leur participation, les participants ont reçu une copie de l'image structurelle anatomique de leur cerveau.

Passation

Préalablement à la session IRM, un premier entretien avait pour but d'expliquer aux participants le déroulement général de l'étude, les différentes tâches perceptives et motrices ainsi que de procéder à un enregistrement des voyelles de chaque sujet devant servir de stimuli auditifs pour la tâche de perception de parole (voir la partie Méthode du Chapitre 1.2).

Lors de l'étude, toutes les étapes et tâches à réaliser ont été rappelées aux participants, une courte session d'entraînement permettant de vérifier leur compréhension et réalisation correcte. Les participants étaient ensuite installés dans le lit du scanner et instruits de la possibilité d'abandonner l'étude à tout moment. Afin de minimiser de possibles artefacts dus aux mouvements de tête et d'éviter des activations cérébrales motrices correspondantes à d'autres parties du corps que les

articulateurs orofaciaux, il était demandé aux participants de rester immobile pendant toute la session IRM.

Un système casque/microphone compatible IRM permettait l'écoute et l'enregistrement des voyelles du participant et, d'autre part, de communiquer avec lui en dehors des runs fonctionnels et anatomique. Les participants portaient également des bouchons d'oreille et un casque antibruit. Les consignes visuelles étaient projetées grâce au logiciel Presentation (Neurobehavioral Systems, Albany, EU) et au moyen d'un vidéo projecteur sur un écran situé derrière le participant et, par réflexion, sur un miroir placé au-dessus de leurs yeux.

Dans le but d'ajuster l'intensité des stimuli sonores perçus lors de la tâche de perception des voyelles avec le volume sonore du feedback auditif lors de la tâche de production des voyelles (malgré les mécanismes de conduction osseuse), une procédure de réglage sonore a été effectuée individuellement pour chacun des participants (voir Christoffels, Formisano et Schiller, 2007). Pour ce faire, le sujet devait dans un premier temps produire ouvertement une série de voyelles de manière la plus naturelle possible. Suite à cette production, l'écoute d'une série de voyelles propres au sujet et préalablement enregistrées permettait au sujet une comparaison du volume sonore des voyelles produites et perçues et, si nécessaire, une modification du volume d'écoute.

Protocole

La Figure 1.1 illustre le déroulement général d'une session IRM pour chacun des deux groupes participant aux études présentées dans les deux premières parties de cette thèse. Chaque session est composée des étapes suivantes :

1. 3 runs fonctionnels relatifs à la tâche de perception de voyelles (3 x 10.5 minutes)
2. 1 run anatomique (15 minutes)
3. 3 runs fonctionnels relatifs à la tâche de production de voyelles (3 x 10.5 minutes)
4. 1 run fonctionnel relatif à la tâche de production de gestes articulatoires supralaryngés (12.5 minutes)



Figure 1.1 : Représentation de l'ordre des acquisitions fonctionnelles et anatomique lors d'une session IRM.

Les images fonctionnelles relatives à l'étude des corrélats neuronaux des systèmes de perception et production des voyelles étaient issues des étapes 1 (perception de voyelles) et 3 (production de voyelles). Celles relatives à l'étude des corrélats neuronaux des gestes supralaryngés et laryngés étaient issues des étapes 3 et 4 (production des gestes supralaryngés). Le scan anatomique acquis lors de l'étape 2 servait de référence anatomique pour la procédure de recalage et normalisation des images fonctionnelles dans un même espace pour les deux études.

L'ordre de passation des différentes étapes était fixe, la tâche de perception auditive précédant celle de production des voyelles en vue de minimiser d'éventuels effets de simulation motrice lors de l'écoute des voyelles.

L'étude des corrélats neuronaux des systèmes de perception et production des voyelles consistait ainsi en 6 runs fonctionnels chacun composé de 63 essais et d'une durée de 10,5 minutes. Lors de chaque run, les 9 voyelles /, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/ ont été perçues ou produites 6 fois selon un ordre pseudoaléatoire (de manière à ce qu'une même voyelle ou un même trait phonétique d'arrondissement, d'antériorité ou d'ouverture ne soit pas jamais répété). Pour chaque run, une condition de repos, devant servir de ligne de base, était également présentée 6 fois. Du fait de

l'utilisation d'une procédure de type "sparse sampling" pour l'acquisition des images fonctionnelles (voir ci-dessous), l'ordre des délais d'acquisition de ces images a également été contrebalancé (de manière à ce que le même délai d'acquisition ne se répète jamais lors de deux scans fonctionnels successifs). Enfin, 3 scans fonctionnels ont été ajoutés en début de chaque run de manière à équilibrer le signal IRM et ont ensuite été supprimés des analyses. En tout, 360 scans ont ainsi été acquis incluant pour chacune des tâches de perception et de production des voyelles 18 images fonctionnelles par voyelle et 18 images fonctionnelles pour la condition de repos.

L'étude des corrélats neuronaux du système de contrôle moteur des gestes supralaryngés et laryngés consistait en la réalisation des 3 tâches motrices supralaryngées (protrusion des lèvres, rétraction horizontale de la langue, ouverture mandibulaire) lors du dernier run fonctionnel et de la production de la voyelle /i/ réalisée lors des 3 runs fonctionnels précédents³³. Chaque condition consistait en 18 essais présentés selon un ordre pseudoaléatoire (aucune répétition successive de la même condition). Comme lors des runs relatifs à la production de voyelles, une condition de repos était également présentée 18 fois lors du run fonctionnel lié au gestes supralaryngés. L'ordre des délais d'acquisition des images était également contrebalancé et 3 scans fonctionnels ont été ajoutés en début de run. En tout, 108 scans ont ainsi été acquis incluant pour chacune des conditions motrices 18 images fonctionnelles par voyelle ainsi que 18 images fonctionnelles pour les deux conditions de repos relatives à la tâche de production des voyelles et celle de production des gestes supralaryngés.

A l'exception de l'utilisation d'un paradigme d'adaptation, le paradigme expérimental des études IRMf présentées dans la seconde partie de cette thèse était en tous points identique. De manière à caractériser un possible effet d'adaptation (répétition-suppression), chaque condition était néanmoins réalisée par train consécutifs de 6 essais d'une même condition selon un ordre pseudoaléatoire (consistant en la perception ou production répétée 6 fois d'une même voyelle ou de la réalisation répétée 6 fois d'un même geste articulatoire supralaryngé).

Acquisition des données IRM

Les images anatomique et fonctionnelles ont été acquises avec un imageur corps entier à 3 Tesla (Bruker Medspec S300) muni d'une antenne tête émission/réception avec champ de vue large. Pour l'acquisition des scans fonctionnels, une séquence d'acquisition de type EPI (Echo Planar Imaging) en écho de gradient pondérée en T2* a été utilisée. Une séquence de repérage initial a permis l'acquisition de volumes fonctionnels orientés parallèlement au plan sagittal et alignés selon les commissures antérieure et postérieure (plan CA-CP). Pour chaque volume fonctionnel, quarante coupes axiales adjacentes ont ainsi été acquises en mode entrelacé. Les paramètres des séquences EPI étaient les suivants : temps de répétition (TR) de 10s, temps d'acquisition (TA) de 2600ms, temps d'écho (TE) de 30ms, champ de vue de 216 mm², matrice d'acquisition de 72x72 pixels, matrice de reconstruction de 128*128 pixels, taille des voxels de 3 mm³ sans interstice, angle de bascule de 90°. Afin de permettre une normalisation spatiale des données fonctionnelles sur un modèle de référence, une image anatomique de haute résolution (séquence pondérée en T1 de type MP-RAGE) a également été acquise pour chaque participant (volume de 256x224x176mm, résolution isotrope de 1mm³, angle de bascule de 30°).

³³ L'étude IRMf portant sur les mécanismes d'adaptation sensorimotrice des gestes orofaciaux présentée dans la seconde partie de cette thèse ne porte que sur les seuls gestes supralaryngés (mouvement des lèvres, de la langue et de la mandibule) et n'implique que le dernier run fonctionnel.

Méthode d'acquisition sparse sampling

Afin de minimiser de possibles artefacts sur les images fonctionnelles, liés à une déformation du champ magnétique du fait de mouvements orofaciaux et d'une modification du conduit vocal, ainsi que de limiter le bruit du scanner inhérent à l'acquisition des images, un paradigme de type sparse sampling (Birn et al., 1999; Hall et al., 1999; Gracco, Tremblay et Pike, 2005) a été utilisé dans toutes les études IMRF présentées dans les deux premières parties de cette thèse. La méthode d'acquisition en sparse sampling est basée sur le délai existant entre l'activité neuronale liée à une tâche motrice ou à l'écoute d'un stimulus auditif et le pic maximal de la réponse hémodynamique associée. Face au délai optimal estimé à 5s dans de précédentes études du pic de la réponse hémodynamique lors de la production de mouvement orofaciaux ou lors de la perception auditive ou production de séquences de parole (Gracco, Tremblay et Pike, 2005; Özdemir, Norton et Schlaug, 2006; Sörös et al., 2006; Zaehle et al., 2007), l'intervalle de temps séparant la perception d'une voyelle, la production d'une voyelle ou la réalisation d'un geste orofacial et le milieu d'acquisition de l'image fonctionnelle correspondante était contrebalancé entre chaque essai et était de 4s, 5s ou 6s (voir Figure 1.2).

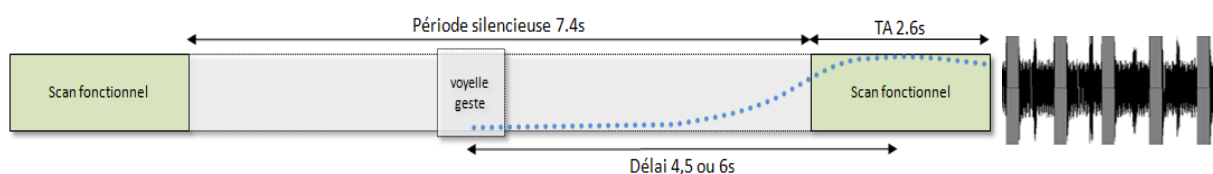


Figure 1.2 : Gauche : Représentation d'un essai lors de l'acquisition d'une image fonctionnelle. Les scans fonctionnels sont acquis durant 2.6s toutes les 10s. La perception et production des voyelles ou des gestes supralaryngés est réalisée lors d'une période silencieuse du scanner, 4s, 5s ou 6s avant l'acquisition de l'image fonctionnelle correspondante. La courbe en pointillé représente le déroulement temporel schématisé de l'amplitude de la réponse hémodynamique évoquée par la perception d'un stimulus auditif ou par la production d'un geste vocalique ou supralaryngé. Droite : Enregistrement acoustique de 4 trials successifs lors d'une tâche de production de voyelles. Les barres grises représentent le bruit du scanner lors de l'acquisition d'une image fonctionnelle (2.6s), la ligne noire épaisse dans le sens horizontal représente le bruit stationnaire de la pompe à hélium lors de la période d'inactivité du scanner (7.4s), les barres noires fines représentent la production maintenue d'une voyelle (de l'ordre de 500ms).

Lors des tâches de production de voyelles et de gestes orofaciaux, une consigne visuelle indiquait aux participants la condition motrice à réaliser (type de voyelle ou de mouvement orofacial supralaryngé) et était affichée 4.5s, 5.5s ou 6.5s avant le milieu d'acquisition du scan fonctionnel correspondant. Par comparaison avec l'écoute des stimuli auditifs lors de la tâche de perception des voyelles, ce délai supplémentaire de 500ms avait pour but la prise en compte de la lecture de la consigne visuelle et l'initiation du mouvement et correspondait précisément à celui observé lors de l'enregistrement des voyelles par les participants effectué préalablement à la passation IRM (voir la partie Méthode du Chapitre 1.2).

Le choix de ces 3 délais d'acquisition, décidé en fonction des résultats d'études précédentes, devait permettre une capture optimale de la réponse hémodynamique pour tous les participants, ce malgré des variations supposées d'activations à la fois entre individus et également entre régions. En effet, les paramètres du signal BOLD dépendent non seulement de l'activité neuronale mais aussi de la distance entre la région activée et l'artère qui l'irrigue, tout comme le diamètre de celle-ci (Logothetis et al. 2001). De plus, l'utilisation de ces 3 délais permettait de contrôler *a posteriori* de possibles variations du déroulement temporel du signal hémodynamique en fonction de la tâche perceptive ou motrice réalisée et, de là, de fournir des informations potentiellement importantes pour de futures études quant à l'optimisation des paramètres d'acquisition et une capture optimale des réponses hémodynamiques des régions cérébrales spécifiquement associées à la perception et production de la parole et à la production de mouvements orofaciaux.

Prétraitements

Les données recueillies depuis la plateforme IRM sous format Bruker ont subi une première conversion sous format DICOM (Digital Imaging and Communications in Medicine, format standard permettant le transfert d'images entre machines de différents constructeurs) puis sous le format NIfTI (Neuroinformatics Technology Initiative's Data Format) exploité par le logiciel SPM5 (pour Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK).

Préalablement aux analyses statistiques, une série de prétraitements des images fonctionnelles a été réalisée avec le logiciel SPM5 sous environnement Matlab 7.1 (Mathworks, Natick, MA). Ces prétraitements incluaient des procédures de réalignement, de normalisation dans l'espace commun du Montreal Neurological Institute (repère MNI) et de lissage spatial (voir Figure 1.3).

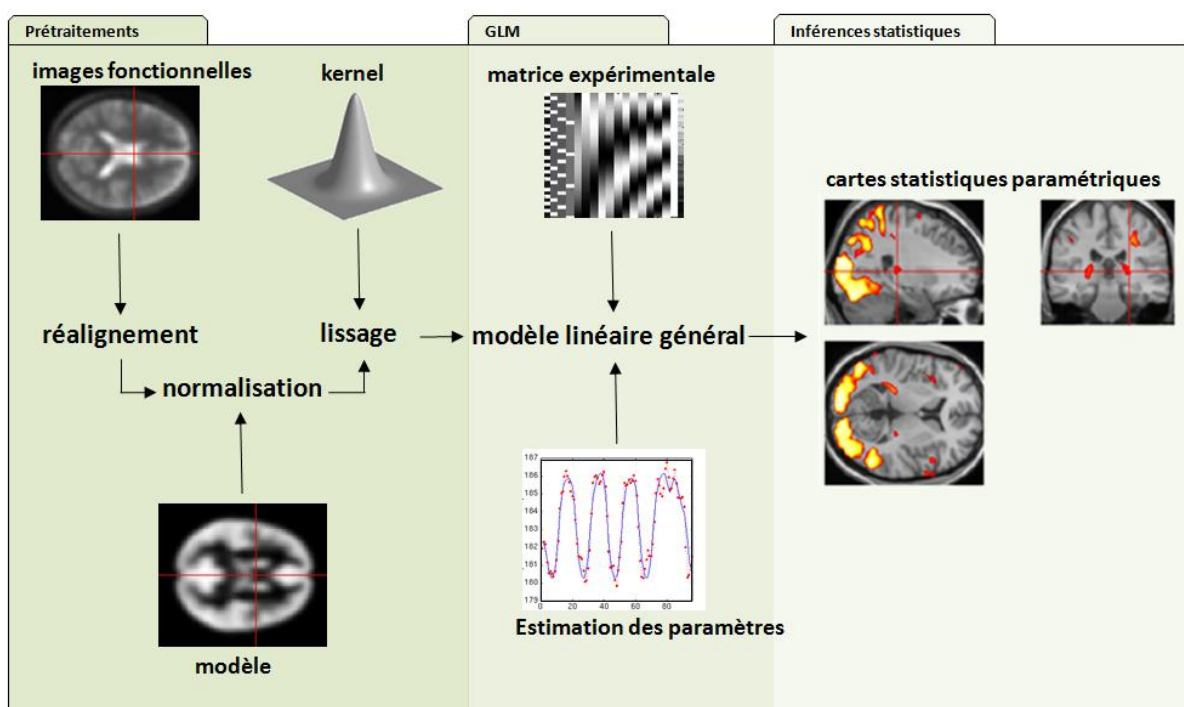


Figure 1.3 : Etapes successives nécessaires aux prétraitements et à l'analyse des images fonctionnelles. L'établissement de cartes statistiques paramétriques fait suite à des procédures de réalignement, de normalisation dans un espace commun, de lissage spatial, de la définition d'un modèle linéaire général fonction du protocole expérimental et de l'estimation des paramètres de ce modèle.

Slice Timing

Il est généralement conseillé d'effectuer un recalage temporel (slice timing) entre images fonctionnelles du fait du décalage d'acquisition des différentes coupes de chaque image. Cette procédure étant basée sur des interpolations temporelles entre coupes et images successives, ce prétraitement n'a cependant pu être réalisé du fait de l'utilisation d'une procédure d'acquisition sparse sampling et d'un délai d'acquisition entre images de 10s (Bohland et Guenther, 2006).

Réalignement des images fonctionnelles

Les images fonctionnelles de chacun des participants ont été réalignées selon une procédure de transformation/interpolation géométrique affine du déplacement rigide de chaque image par

rapport à une image de référence (moyenne de la première image fonctionnelle de chaque run). Avant le réalignement des images, ces paramètres de mouvement (translation en mm et rotations en degré sur/autour des axes médio-latéral x, rostro-caudal y et dorso-ventral z) étaient contrôlés pour chaque sujet. Aucun déplacement supérieur à 3mm et/ou 3° par rapport à la première image fonctionnelle de chaque run n'a été constaté (voir Figure 1.4).

Recalage anatomo-fonctionnel (coregister)

Une procédure de recalage anatomo-fonctionnel a été réalisée de manière à établir une correspondance entre l'image anatomique de haute définition (T1) et l'image moyenne des images fonctionnelles (EPI) et de permettre une réorientation de l'image anatomique par des procédures de normalisation de l'information mutuelle et d'interpolation géométrique.

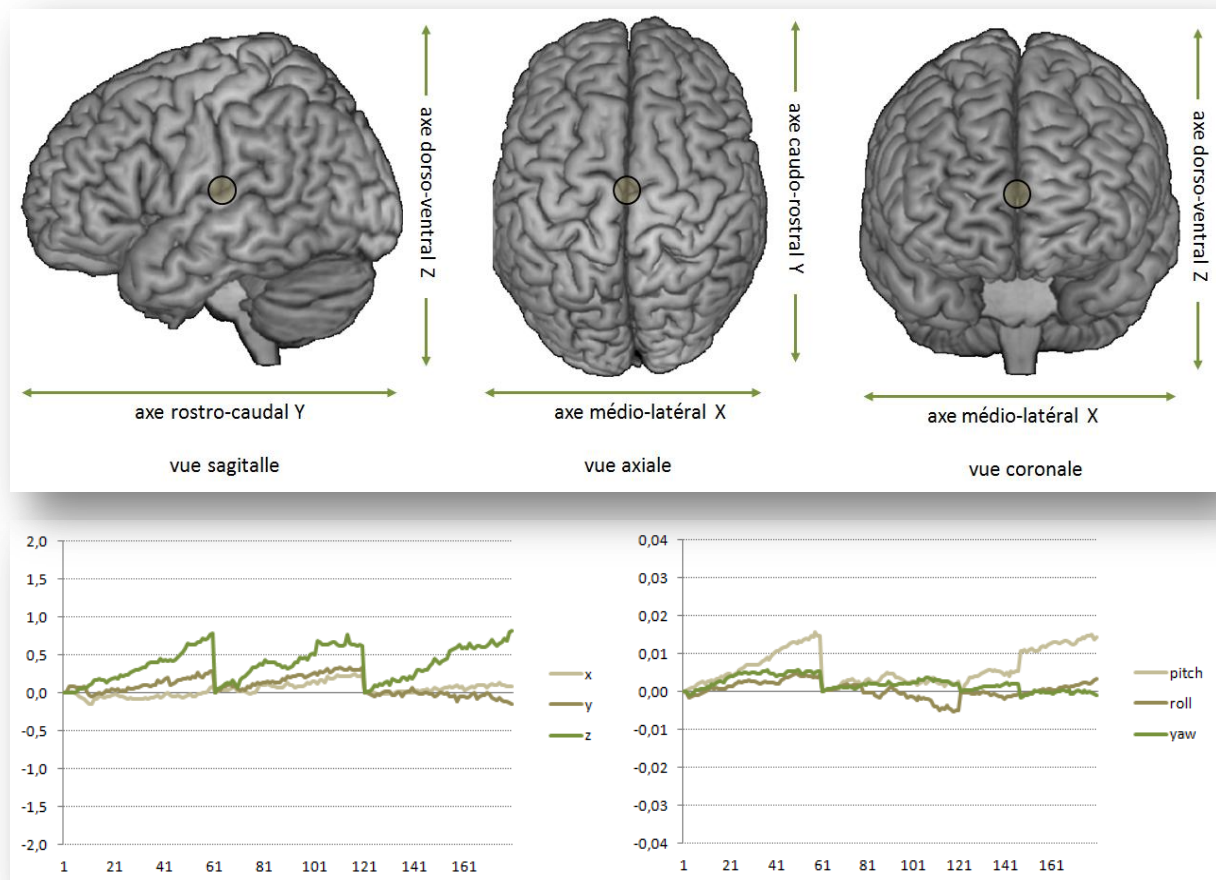


Figure 1.4. Haut : Système de référence anatomique, axes médio-latéral x, rostro-caudal y et dorso-ventral z, le cercle représente l'origine du système. Bas : Paramètres de mouvement pour un sujet lors des tâches de perception et de production des voyelles et de réalisation des gestes supralaryngés, paramètres de translation (en mm) et de rotation (en degré) par rapport aux axes médio-latéral x, rostro-caudal y et dorso-ventral z.

Segmentation

Une procédure de segmentation de l'image anatomique (T1) en différents tissus (matière grise, matière blanche, liquide céphalo-rachidien) a ensuite été réalisée en regard des templates implémentés dans SPM5. Cette procédure permettait également d'estimer les paramètres de réalignement et normalisation nécessaires entre l'image anatomique du sujet et l'espace MNI.

Normalisation spatiale

Par rapport aux différences anatomiques inter-individuelles, il était nécessaire d'effectuer une mise en correspondance spatiale des images fonctionnelles de tous les sujets dans le but de permettre le calcul des coordonnées des régions activées dans un même espace de référence. Il s'agit ainsi de projeter l'ensemble des volumes fonctionnels des sujets dans un espace tridimensionnel commun, l'espace MNI. Du fait de la plus haute résolution spatiale de l'image anatomique, les paramètres de réaligement et de normalisation de l'image anatomique préalablement calculés par rapport à l'espace MNI ont donc été utilisés pour la normalisation des images fonctionnelles par une procédure d'interpolation.

Lissage spatial

Afin de minimiser les variabilités spatiales inter-individuelles résiduelles, d'une part, et d'augmenter le rapport signal sur bruit de manière générale, un filtre gaussien passe-bas de 6 mm^3 a été appliqué sur les images fonctionnelles normalisées.

Chapitre 1.1

Corrélat neuronal des gestes orofaciaux

Résumé / Abstract

Comparé à des actions orofaciales coordonnées et complexes, peu d'études ont analysé les substrats neuronaux communs et distincts de mouvements simples articulatoires supralaryngés et laryngés exécutés de manière indépendante. Afin de déterminer les régions corticales et sous-corticales associées au contrôle moteur supralaryngé, des enregistrements par imagerie par résonance magnétique fonctionnelle ont été réalisés lors de l'exécution de mouvements labiaux, linguaux ou mandibulaires par les participants. Pour les activités laryngées, les participants ont produit la voyelle /i/ de manière maintenue. Une méthode d'acquisition en sparse sampling a été utilisée afin de minimiser les artefacts liés aux mouvements orofaciaux. Trois résultats majeurs ont été observés. Premièrement, les quatre tâches ont activé un ensemble de régions cérébrales communes : les cortex sensorimoteur et prémoteur, le gyrus frontal inférieur droit, l'aire motrice supplémentaire, l'opercule pariétal gauche et le lobule pariétal adjacent, les ganglions de la base et le cervelet. De plus, les différences d'activité entre tâches ont été restreintes au cortex auditif de manière bilatérale et au cortex sensorimoteur ventro-latéral gauche, avec une intensité du signal plus ample pour la tâche de vocalisation. Enfin, une organisation somatotopique dorso-ventrale des mouvements labiaux, mandibulaires, vocaliques/laryngés et linguaux a été observée au sein des cortex moteur et somatosensoriel primaire en utilisant des analyses individuelles par région d'intérêt. Ces résultats démontrent l'existence d'un réseau cortical minimal impliqué dans le contrôle moteur laryngé et supralaryngé et précisent l'organisation somatotopique sensorimotrice des articulateurs orofaciaux.

Compared to complex coordinated orofacial actions, few studies have attempted to determine the shared and distinct neural substrates of supralaryngeal and laryngeal articulatory movements when performed independently. In order to determine cortical and subcortical regions associated with supralaryngeal motor control, participants produced lip, tongue or jaw movements while undergoing functional magnetic resonance imaging. For laryngeal activity, participants produced the steady-state /i/ vowel. A sparse temporal sampling acquisition method was used to minimize movement-related artifacts. Three main findings were observed. First, the four tasks activated a set of largely overlapping, common brain areas: the sensorimotor and premotor cortices, the right inferior frontal gyrus, the supplementary motor area, the left parietal operculum and the adjacent inferior parietal lobule, the basal ganglia and the cerebellum. Second, differences between tasks were restricted to the bilateral auditory cortices and to the left ventrolateral sensorimotor cortex, with greater signal intensity for vowel vocalization. Finally, a dorso-ventral somatotopic organization of lip, jaw, vocalic/laryngeal and tongue movements was observed within the primary motor and somatosensory cortices using individual ROI analyses. These results provide evidence for a core neural network involved in laryngeal and supralaryngeal motor control and further refine the sensorimotor somatotopic organization of orofacial articulators.

- *Mots clés : Contrôle moteur orofacial, somatotopie, production de voyelle, IRMf, sparse sampling.*
- *Keywords: Orofacial motor control, somatotopy, vowel production, fMRI, sparse sampling.*

1.1.1 Introduction

Corrélat neurologiques du contrôle moteur orofacial

Les corrélats neurologiques du contrôle moteur orofacial ont été par le passé très largement étudiés dans le cadre d'actions coordonnées impliquant des mécanismes de co-articulation entre articulateurs supralaryngés et laryngés, comme lors de la mastication (Nakamura et Katakura, 1995; Lund et Kolat, 2006; Onozuka et al., 2007), du déglutissement (Hamdy et al., 1999; Martin et al., 2001, 2004, 2007; Sawczuk et Mosier, 2001; Humbert et Robbins, 2007; Leopold et Daniels, 2010; Peeva et al., 2009), du sifflement (Dresel et al., 2005), de la phonation et de la vocalisation (Jürgens, 2002, 2009; Schulz et al., 2005; Smotherman, 2007; Brown, Ngan et Liotti, 2008; Simonyan et al., 2009) et enfin de la production de la parole (Murphy et al., 1997; Wise et al., 1999; Lotze et al., 2000a; Riecker et al., 2000a,b, 2005, 2008; Brown et al., 2005; Guenther, Ghosh et Tourville, 2006; Bohland et Guenther, 2006; Özdemir, Norton et Schlaug, 2006; Sörös et al., 2006; Terumitsu, et al., 2006; Brown, Ngan et Liotti, 2008; Chang et al., 2009).

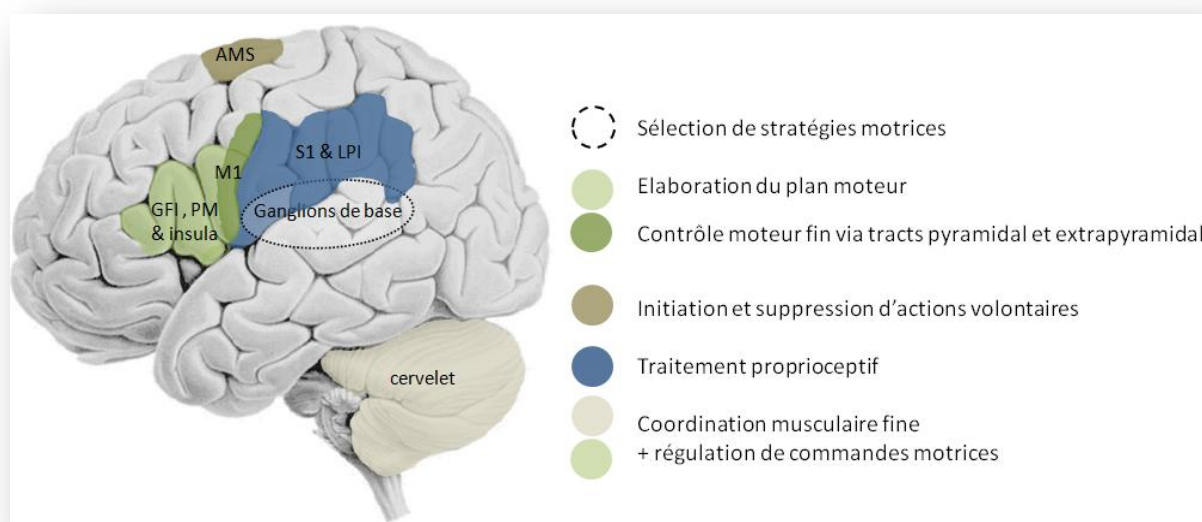


Figure 1.5 : Représentation schématique des régions corticales et sous-corticales impliquées dans le contrôle moteur orofacial. AMS : aire motrice supplémentaire; GFI : gyrus frontal inférieur; PM : cortex prémoteur; M1 : cortex moteur primaire; S1 : cortex somatosensoriel primaire; LPI : lobule pariétal inférieur.

Pris ensemble, les résultats de ces études suggèrent l'existence d'un "réseau neural minimal" (core neural network) lié au contrôle moteur orofacial (voir Figure 1.5), incluant le cortex ventral sensorimoteur et le cortex prémoteur ventral/dorsal de manière bilatérale, l'aire motrice supplémentaire, l'opercule pariétal gauche et le lobule pariétal inférieur adjacent, la partie postérieure (operculaire) du gyrus frontal inférieur, les ganglions de la base et le cervelet. Bien que les fonctions spécifiques de ces régions corticales et sous-corticales restent débattues, elles peuvent être globalement assignées à des mécanismes de préparation, de planification, de coordination et d'exécution motrice ainsi qu'à des boucles de régulation sensori-motrice (voir Jürgens, 2002; Riecker et al., 2005 pour revues). L'initiation et la suppression des mouvements volontaires est traditionnellement assignée à l'aire motrice supplémentaire (Pickard et Strick, 2001) alors que le contrôle moteur fin des gestes orofaciaux est effectué par le cortex moteur primaire et les ganglions de la base par les voies pyramidales et extra-pyramidales (Wise et al., 1999; Jürgens, 2002; Riecker et al., 2005). Le cortex moteur reçoit les plans moteurs élaborés depuis le cortex prémoteur (Pickard et Strick, 2001) et le gyrus frontal inférieur postérieur adjacent, ainsi que les informations

proprioceptives depuis les aires somatosensorielles primaires et associatives et le lobule pariétal inférieur (Smith, 1998; Riecker et al., 2005). Sur ce point, il est à noter que l'implication de la partie postérieure du gyrus frontal inférieur et du lobule pariétal inférieur lors de tâches de production de parole et de gestes orofaciaux reste néanmoins débattue (Murphy et al., 1997; Wise et al., 1999; Lotze et al., 2000a; Corfield et al., 1999; Nota et Honda, 2004; Dhanjal et al., 2008; Chang et al., 2009). Enfin, les ganglions de la base jouent un rôle important non seulement dans les processus de sélection mais aussi dans la régulation des commandes motrices par des projections thalamo-motrices (Houk et al., 2007). En parallèle, le cervelet reçoit des informations motrices et sensorielles et est impliqué dans la coordination musculaire fine des mouvements produits (Thach, 1998; Houk et al., 2007).

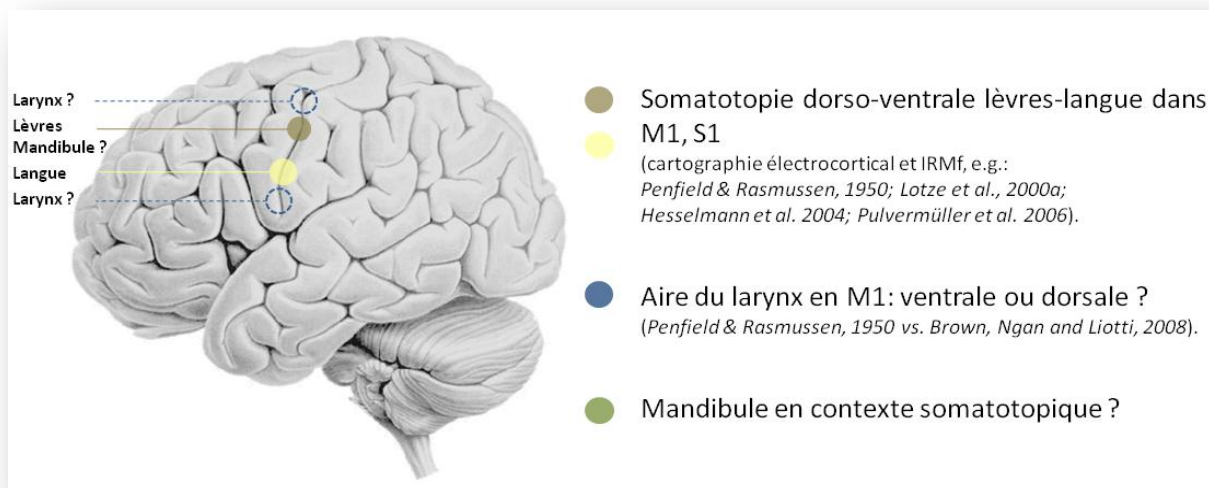


Figure 1.6 : Représentation schématique de l'organisation somatotopique sensorimotrice dorso-ventrale orofaciale.

Par rapport aux actions orofaciales complexes et coordonnées, les substrats neuronaux des mouvements articulatoires supralaryngés et laryngés, lorsqu'exécutés indépendamment, ont cependant été moins étudiés (Lotze et al., 2000a; Hesselmann et al., 2004; Terumitsu et al., 2006; Pulvermüller et al., 2006; Brown, Ngan et Liotti, 2008; Dhanjal et al., 2008; Brown et al., 2009). La plupart des études de neuroimagerie impliquant une production de mouvements orofaciaux indépendants se sont essentiellement concentrées sur une organisation somatotopique possible des articulateurs orofaciaux au sein du cortex sensorimoteur (Lotze et al., 2000a; Hesselmann et al., 2004; Terumitsu et al., 2006; Brown, Ngan et Liotti, 2008; voir aussi Pulvermüller et al., 2006). Les travaux de Wilder Penfield et collaborateurs (Penfield et Boldrey, 1937; Penfield et Rasmussen, 1950) ont été parmi les premiers à démontrer l'existence d'une organisation topographique des cortex moteur et somatosensoriel primaires. Par l'utilisation de stimulations électriques appliquées à des régions distinctes du cortex sensorimoteur, lors d'opérations neurochirurgicales chez des patients éveillés, ils ont en effet observé le déclenchement de mouvements musculaires spécifiques involontaires et ont ainsi pu établir une cartographie corticale liée aux représentations motrices et somesthésiques des différentes parties du corps. Cette organisation corticale sensorimotrice, dite "somatotopique", des parties corporelles ne reflète néanmoins pas la morphologie humaine, l'étendue des zones motrices associées à chaque partie du corps correspondant plutôt à la précision et la sensibilité du contrôle moteur des gestes associés à celle-ci. Malgré des superpositions et chevauchements importants des régions sensorimotrices liées aux gestes orofaciaux, Penfield et Rasmussen (1950) ont pu observer une organisation somatotopique dorso-ventrale au sein du cortex sensorimoteur des lèvres, de la mandibule, de la langue et du pharynx. Ces résultats ont été depuis partiellement reproduits par l'utilisation de techniques d'imagerie non-invasives comme l'IRMf (voir Figure 1.6 et Table 1.2 pour une revue des activités motrices orofaciales précédemment observées).

Ces études ont ainsi confirmé une topographie dorso-ventrale au sein du cortex sensorimoteur des activations liées à des mouvements labiaux et linguaux (Lotze et al., 2000a; Hesselmann et al. 2004; Pulvermüller et al. 2006). L'emplacement d'une aire spécifique liée au contrôle laryngé au sein du cortex moteur primaire apparaît cependant plus contesté, deux études récentes reportant des positions différentes pour les mouvements laryngés et variant également selon l'hémisphère (Terumitsu et al., 2006; Brown, Ngan et Liotti, 2008). Enfin à notre connaissance, aucune étude de neuroimagerie n'a encore tenté de localiser l'aire contrôlant les mouvements mandibulaires et de la replacer dans un contexte somatotopique.

Cadre de l'étude

Alors que l'existence d'une somatotopie sensorimotrice générale ne fait aucun doute à la lumière des précédents travaux, peu d'études ont exploré et directement comparé les corrélats neuronaux des mouvements supralaryngés et laryngés chez les mêmes participants (Brown, Ngan et Liotti, 2008; Dhanjal et al., 2008). Il apparaît pourtant fondamental de préciser plus avant l'organisation des articulateurs orofaciaux au sein des cortex moteur et somatosensoriel, notamment vers une meilleure compréhension des processus sensorimoteurs de production de la parole. La coordination des articulateurs laryngés et supralaryngés lors de la production de la parole apparaît en effet essentielle, aussi bien au niveau segmental que supra-segmental. Dans ce cadre, il est d'ailleurs à noter que certains modèles (neuro)computationnels de production de la parole impliquent des cartes sensorimotrices détaillées dans lesquelles chaque articulateur a un rôle et une représentation spécifique (Guenther, Ghosh et Trouville, 2006; Guenther et Vladusich, 2012). Enfin, aucune étude n'a à ce jour examiné les corrélats neuronaux liés à un mouvement mandibulaire bien que cet articulateur soit fondamental pour la production de la parole et soit considéré comme jouant un rôle crucial dans le développement du contrôle moteur particulièrement lors de l'acquisition de syllabes (MacNeilage, 1998).

Les études précédentes de neuroimagerie portant sur le contrôle moteur orofacial démontrent, malgré de fortes superpositions, une organisation somatotopique générale des articulateurs ("an overall picture of somatotopy with overlap", Takai, Brown et Liotti, 2010) et suggèrent ainsi un système de contrôle sensorimoteur spécifique et intégré pour la production de gestes orofaciaux. Dans la présente étude, nous avons utilisé la méthode d'IRMf dans le but de clarifier les corrélats neuroanatomiques fonctionnels communs et distincts de mouvements simples supralaryngés et laryngés. Les mouvements supralaryngés impliquaient une protrusion des lèvres, un abaissement de la mandibule ou une rétraction de la langue. La production de la voyelle fermée antérieure non arrondie /i/ a été utilisée dans le but de déterminer les activités cérébrales liées au contrôle laryngé. Ce choix, déterminé *a posteriori* et en fonction de l'étude annexe de production des voyelles du Français (voir Chapitre 1.2), reste bien évidemment contestable du fait d'une implication concomitante des articulateurs supralaryngés lors de la production de la voyelle. Il a cependant été montré qu'une région spécifique au sein du cortex moteur primaire et supposée représenter le contrôle du larynx est comparablement activée lors de mouvements laryngés aussi bien vocaliques que non vocaliques (production du schwa vs adduction forcée des cordes vocales en absence de tout voisement, Brown, Ngan et Liotti, 2008). De plus, face à ce choix d'une production vocalique, la vocalisation de la voyelle fermée antérieure non arrondie /i/ implique potentiellement une activité moindre des muscles labiaux et mandibulaires que les autres voyelles arrondies ou plus ouvertes.

Face aux résultats précédemment observés, la présente étude a ainsi pour but de préciser à la fois (1) les substrats neuronaux communs et distincts des mouvements supralaryngés et laryngés et (2) l'organisation somatotopique sensorielle et motrice des mouvements labiaux, mandibulaires, linguaux et vocaliques/laryngés, (3) en utilisant une méthode d'acquisition dite de "sparse sampling" en vue de minimiser de possibles artefacts des images fonctionnelles liés aux mouvements des articulateurs.

1.1.2 Méthode

Protocole

Participants

Treize volontaires adultes sains ont participé à l'étude (11 hommes; âge : 21-44 ans, moyenne : 29 ans). Le recrutement des sujets est explicité dans la partie "Considérations Méthodologiques".

Procédure expérimentale

L'expérience consistait en la réalisation distincte des tâches suivantes :

- Une protrusion des lèvres (condition "lèvres")
- Une rétraction horizontale de la langue (condition "langue")
- Une ouverture mandibulaire (condition "mâchoire")
- Une condition de contrôle laryngé correspondant à la production de la voyelle fermée antérieure non arrondie /i/ (condition "voyelle")

Il était demandé aux participants d'initier chacune de ces tâches depuis une position immobile de repos, les muscles linguaux et labiaux relâchés et de revenir à cette position initiale de repos à la fin du mouvement.

De manière à évaluer les activations cérébrales spécifiquement associées au contrôle moteur de chacun de ces articulateurs, une condition "contrôle", sans aucune activité motrice, a également été testée.

Nous avons décidé *a posteriori* de comparer les corrélats neuronaux des gestes supralaryngés aux gestes laryngés/vocaliques issus de la tâche de production de voyelles effectuée par les mêmes participants lors de la même session expérimentale (voir Chapitre 1.2). Par rapport aux voyelles /i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/ produites lors de cette tâche, nous nous sommes focalisés sur la voyelle fermée antérieure non arrondie /i/ dont la réalisation est supposée requérir une implication et coordination moindre des articulateurs supralaryngés tels les muscles de la mandibule (voyelles fermées), les lèvres (voyelles non arrondies). Bien que la production d'une voyelle ait été utilisée dans de précédentes études dans le but de déterminer les activités neuronales laryngées (Terumitsu et al., 2006; Brown, Ngan et Liotti, 2008), ce choix *a posteriori* est bien sûr contestable, il serait notamment plus judicieux d'utiliser une tâche de vibration des cordes vocales, comme par exemple un bourdonnement sourd, afin d'étudier plus spécifiquement le contrôle des muscles laryngés.

Les 3 tâches motrices supralaryngées ont été réalisées lors d'un même run fonctionnel, la tâche de production de la voyelle /i/ étant réalisée lors de 3 runs fonctionnels précédant celui qui correspond aux mouvements supralaryngés (voir partie "Considérations Méthodologiques", Figure 1.1.). Durant l'expérience, le sujet devait réaliser l'une des 4 conditions motrices ou la condition de repos toutes les 10 secondes (du fait de l'acquisition des images fonctionnelles de type sparse sampling) selon un ordre pseudo-aléatoire (une même condition ne pouvant survenir plus de deux fois de suite). Chaque condition consistait en 18 essais. Pour chaque essai, une consigne visuelle ("lèvres", "langue", "mâchoire", "ii" ou "repos") indiquait au sujet la condition à réaliser. Précédé d'une croix de fixation durant 500ms, l'affichage de cette consigne durant 1s servait également de repère temporel aux sujets pour la réalisation de la tâche. Concernant le protocole d'acquisition en sparse sampling, l'intervalle de temps séparant un geste orofacial et le milieu d'acquisition de l'image fonctionnelle correspondante était contrebalancé entre chaque essai et était de 4s, 5s ou 6s. 108 scans fonctionnels ont ainsi été acquis ((4 conditions motrices + 2 condition de repos) x 18 scans). Lors de

chaque run fonctionnel, 3 scans ont été ajoutés en début de session pour équilibrer le signal IRM et ont ensuite été supprimés des analyses.

Prétraitements des images fonctionnelles

Préalablement aux analyses statistiques, une série de prétraitements des images fonctionnelles a été réalisée avec le logiciel SPM 5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) sous environnement Matlab 7.1 (Mathworks, Natick, MA). Ces prétraitements incluaient des procédures de réalignement, de normalisation dans l'espace commun du Montreal Neurological Institute (repère MNI) et de lissage spatial (voir partie "Considérations Méthodologiques" pour une description détaillée des ces procédures).

Analyses statistiques

Suite aux prétraitements, les données fonctionnelles ont été analysées avec le logiciel SPM 5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) sous environnement Matlab 7.1 (Mathworks, Natick, MA). Les régions cérébrales activées ont été labellisées en utilisant la toolbox SPM Anatomy (Eickhoff et al., 2005) et, si nécessaire, en utilisant le logiciel Talairach Daemon (Lancaster et al., 2000). Les analyses par région d'intérêt (ROI) ont été effectuées via la toolbox SPM Anatomy. Pour la visualisation des résultats, les cartes d'activation ont été superposées sur un template anatomique standard en utilisant le logiciel MRICRON (<http://www.sph.sc.edu/comd/rorden/mricron/>).

Analyses individuelles

Au niveau individuel et suite aux prétraitements des images fonctionnelles (voir partie "Considérations Méthodologiques"), les activations neuronales liées aux gestes orofaciaux supralaryngés et vocaliques ont été estimées par l'utilisation d'un Modèle Linéaire Général (GLM; Friston et al., 1995). Pour chaque participant, un second GLM a été estimé de manière à évaluer les activités reliées aux 3 délais d'acquisition (4s, 5s et 6s) indépendamment des tâches motrices.

Les modèles incluaient des régresseurs d'intérêt reliés soit aux 4 tâches motrices (chacune représentée par 18 images fonctionnelles) indépendamment des délais d'acquisition, soit aux 3 délais d'acquisition indépendamment des tâches motrices (chacune représentée par 18 images fonctionnelles issues du run fonctionnel lié aux mouvements supralaryngés) et des régresseurs de non-intérêt liés aux paramètres de réalignement. Les conditions de repos formaient une ligne de base.

Pour les deux modèles, la réponse de type hémodynamique associée à chaque évènement a été modélisée par une fonction à réponse impulsionnelle finie et à impulsion unique (FIR) pour chaque scan fonctionnel. Avant l'estimation du modèle, un filtrage des basses fréquences *a priori* non-relies aux conditions expérimentales (variations lentes d'origine physiologique) a été appliqué (passe-haut de fréquence de coupure de 1/128 Hz). Les estimations des réponses FIR modélisées ont été calculées de manière à correspondre avec le dérouls temporel du signal BOLD observé pour chaque voxel et chaque condition. Pour chaque participant, des cartes d'analyse statistique individuelles ont ainsi été calculées pour chacune des quatre tâches motrices dans le premier modèle et pour chacun des 3 délais d'acquisition dans le second modèle, par contraste avec la condition contrôle. Ces cartes statistiques individuelles ont été ensuite utilisées pour les analyses de groupe.

Analyses de groupe

Afin de permettre une généralisation des résultats observés pour la population étudiée (Friston, Holmes et Worsley, 1999), des analyses de groupe de second niveau à effet aléatoire ont été réalisées au moyen d'analyses de variance (ANOVA) à mesures répétées.

Mouvements orofaciaux

De manière à déterminer les activités cérébrales liées aux mouvements orofaciaux, une ANOVA à mesures répétées a été réalisée avec comme facteur intra-sujet la tâche motrice (incluant 4 niveaux liés aux mouvements vocaliques, des lèvres, de la langue et de la mandibule). La variable "sujets" a été prise en compte comme facteur aléatoire. Du fait de la répartition des tâches motrices dans différents runs fonctionnels, les 4 conditions motrices ont été définies comme ayant des variances d'activité non homogènes de manière à ajuster les statistiques et les degrés de liberté dans le calcul des inférences statistiques.

Quatre contrastes t unidirectionnels ont été calculés par des tests de Student afin de déterminer les régions cérébrales activées de manière spécifique pour chacune des tâches orofaciales par rapport à la condition de repos. Dans le but d'identifier les activations communes/partagées par toutes les conditions motrices, une analyse de conjonction (Friston et al., 1999; Nichols et al., 2005) a également été effectuée. Finalement, un contraste F bidirectionnel correspondant à l'effet principal de la condition motrice a été calculé dans le but de déterminer les différences d'activations cérébrales entre les quatre tâches motrices.

Délais d'acquisition

Une seconde analyse de groupe sur les données issues du run fonctionnel lié aux gestes supralaryngés a été réalisée dans le but de comparer les activations cérébrales fonctions du délai d'acquisition des images fonctionnelles, indépendamment des tâches motrices supralaryngées. En plus du décours temporel des activations, cette analyse devait ainsi permettre de vérifier l'adéquation des délais d'acquisition choisis. Pour ce faire, une seconde ANOVA à mesures répétées a été effectuée avec comme facteur intra-sujet le délai d'acquisition (incluant 3 niveaux liés aux délais d'acquisition de 4s, 5s et 6s par rapport à la réalisation de la tâche motrice). La variable "sujets" a été prise en compte comme facteur aléatoire.

Trois contrastes t ont été calculés de manière à déterminer les régions cérébrales activées de manière spécifique pour chacun des trois délais d'acquisition par rapport à la ligne de base. Dans le but d'identifier des activations communes et distinctes aux 3 délais d'acquisition, des analyses de conjonction et d'effet principal du délai d'acquisition ont été calculées.

Somatotopie sensorimotrice

Comme une somatotopie sensorimotrice précise des articulateurs orofaciaux n'a pas pu être mise en évidence par l'analyse de groupe, des analyses par régions d'intérêt (ROI) ont été effectuées afin de tester une possible organisation somatotopique des activations supralaryngées et laryngées au sein des structures primaires motrice et somatosensorielle. Trois régions d'intérêt spécifiques correspondant aux cortex primaire moteur et somatosensoriel ont été créés en utilisant la toolbox SPM Anatomy (Eickhoff et al., 2005; voir Figure 1.9). La première ROI correspondait au cortex moteur primaire et était composée des cartes cytoarchitectoniques des aires de Brodmann BA4a et BA4p (parties antérieure et postérieure du cortex moteur primaire; voir Geyer et al., 1996). Etant données les positions anatomiques des régions du cortex somatosensoriel primaire BA3 et BA1 (partie intérieure du sulcus central et partie postérieure du gyrus postcentral respectivement; voir Geyer et al., 1999), deux ROIs indépendantes ont été créées pour analyser les activités du cortex somatosensoriel primaire correspondant aux cartes cytoarchitectoniques des aires 3a et 3b

combinées pour correspondre à l'aire de Brodmann BA3 et à la carte cytoarchitectonique de l'aire BA1³⁴.

Les analyses par ROI ont été réalisées individuellement pour chaque participant et chaque tâche, en appliquant une correction de type "small volume correction" pour chaque ROI selon un seuil statistique défini à $p < .0001$ non corrigé. Les coordonnées MNI du pic d'activation maximale et du centre de gravité (COG) au sein des cortex primaire moteur et somatosensoriel (BA4, BA3 et BA1) ont été déterminées dans les axes médio-latéral (X), rostro-caudal (Y) et dorso-ventral (Z) et ce pour chaque participant, chaque hémisphère et pour toutes les conditions motrices. Les données de deux sujets ont été écartées de cette analyse, car le seuil choisi n'a pas permis de faire apparaître des activations sensorimotrices pour toutes les tâches et les deux hémisphères.

Pour chaque ROI, des ANOVA à mesures répétées à deux facteurs ont été calculées. Les facteurs intra-sujet impliquaient l'hémisphère (2 niveaux : gauche, droite) et l'articulateur (4 niveaux : lèvres, langue, mandibule, larynx). Les analyses ont été effectuées pour les dimensions x, y et z indépendamment (en valeurs absolues des coordonnées MNI pour l'axe x). Pour ces analyses le niveau de significativité a été défini à $p < .05$ et une correction Greenhouse-Geisser a été appliquée en cas de non sphéricité des données. Lorsque nécessaire, des analyses post-hoc ont été effectuées en utilisant des tests de type LSD Fischer.

1.1.3 Résultats

Les résultats des analyses de groupe liées aux réseaux neuronaux des mouvements orofaciaux et des délais d'acquisition sont reportés selon un seuil statistique défini à $p < .05$ après correction de type "Family-Wise Error" (FWE; Nichols et Hayasaka, 2003) et une taille minimale des clusters de 10 voxels. Pour tous les contrastes, les pics d'activation maximum ont été déterminés dans chaque cluster, leur localisation a ensuite été labellisée avec la boîte à outils Anatomy³⁵ de SPM (Eickhoff et al., 2005). Si une région n'avait pas pu être assignée avec Anatomy, elle a été déterminée avec le logiciel Talairach Daemon (Lancaster et al., 2000) grâce aux coordonnées du pic d'activation converties de l'espace MNI à l'espace stéréotaxique standard de Talairach et Tournoux (1988).

Réseaux neuronaux des mouvements orofaciaux

Réseau articuloire commun

Les résultats de l'analyse de conjonction montrent l'activation de régions corticales et sous-corticales largement communes pour les 4 tâches motrices et classiquement impliquées dans le contrôle

³⁴ L'aire BA2 n'a pas été prise en compte dans le contexte de cette étude s'intéressant aux somatotopies orofaciales au sein des structures sensorimotrices primaires. Elle n'est pas unanimement considérée comme aire somatosensorielle primaire (voir par exemple le chapitre "Somatosensory system" de Jon H. Kaas dans "The Human nervous system", 2004). En effet, les subdivisions du cortex somatosensoriel primaire selon Brodmann (1909) et Vogt et Vogt (1919, 1926) comprennent 4 sous-régions : aires 1, 2, 3a et 3b. Tandis que l'aire 3a traite les informations proprioceptives, les aires 3b et 1 traitent les informations tactiles (de bas et haut niveau, respectivement), l'aire 2 serait impliquée dans les traitements de stimuli plus complexes (reconnaissance de formes et de tailles) et dans les stades précoces du contrôle moteur. De plus, il a été montré que chez le macaque les représentations somatotopiques dans l'aire 2 seraient multiples et plus complexes que dans les autres aires du S1 (Pons et al., 1985b).

³⁵ La Toolbox SPM Anatomy intègre des cartes cytoarchitectoniques probabilistes établies d'après des études post-mortem et inclut notamment l'amygdale, l'hippocampe, des aires du cortex visuel primaire et extrastrié (V1, V2, V5/MT), les régions operculaire et triangulaire du gyrus frontal inférieur (44, 45), le cortex prémoteur (6), les cortex moteur et somatosensoriel primaires (4a/4p, 1, 2, 3a, 3b), le cortex auditif primaire et des régions du gyrus temporal supérieur (TE1, TE2, TE3), le lobule pariétal inférieur et les opercules pariétaux (OP1-4), le sillon intra-pariétal et le lobule pariétal supérieur, l'insula et le cervelet.

moteur orofacial. La Table 1.1 répertorie les activations principales pour chacune des conditions motrices (par rapport à la ligne de base), tout comme les activations communes révélées par l'analyse de conjonction. Les cartes d'activations sont illustrées dans la Figure 1.8. Ce réseau minimal des mouvements supralaryngés et laryngés implique, de manière bilatérale, des activations au niveau du sillon central s'étendant rostralement au gyrus précentral et caudalement au gyrus postcentral, avec deux larges clusters bilatéraux activés comprenant les portions supérieures des cortex prémoteur ventral, moteur et somatosensoriel primaires. De plus, deux clusters bilatéraux plus petits au sein du gyrus précentral ont été repérés au niveau du cortex prémoteur dorsal. Des activations bilatérales ont également été trouvées dans l'aire motrice supplémentaire (SMA) et les régions supérieures du cervelet (région "declive" du néocerevet). Des activations frontales sont observées à la frontière des gyri frontaux inférieur et moyen dans l'hémisphère droit (partie operculaire du gyrus frontal inférieur). Enfin, des activations additionnelles sont observées dans l'hémisphère gauche dans le striatum dorsal des ganglions de la base (putamen) et autour de la région ventrale du gyrus postcentral, incluant l'opercule pariétal et s'étendant au gyrus supramarginal.

Activations motrices fonction de l'articulateur

En dehors du réseau minimal des tous les articulateurs et des différences observées dans l'analyse l'effet principal, des activations supplémentaires ont été reportées pour chacun des gestes orofaciaux (voir Table 1.1 et Figure 1.8). Ces activations n'apparaissant ni dans l'analyse de conjonction ni dans l'analyse de l'effet principal, ces résultats sont cependant à considérer avec précaution.

- **Lèvres** : Des activations bilatérales propres à la tâche de protrusion de lèvres ont été trouvées au niveau du gyrus cingulaire moyen et du lobule pariétal supérieur ventral. Des activations dans l'hémisphère gauche sont également observées au niveau du gyrus frontal supérieur, de l'opercule pariétal, du cortex insulaire, des ganglions de la base (globus pallidus) et du thalamus.
- **Mandibule** : Pour la tâche d'abaissement mandibulaire, des activations sont également observées au niveau du gyrus cingulaire moyen, du striatum dorsal droit des ganglions de la base (putamen), de l'opercule pariétal gauche et du cortex insulaire adjacent et, enfin, dans le gyrus temporal transverse gauche.
- **Langue** : En ce qui concerne les mouvements linguaux, des activations supplémentaires ont été trouvées bilatéralement au sein du lobule pariétal supérieur et du thalamus et, dans l'hémisphère droit, au sein du cortex insulaire antérieur et des ganglions de la base (globus pallidus et striatum dorsal incluant le putamen et le noyau caudé). De manière générale, il est également à constater une étendue des régions activées au sein du cortex sensorimoteur plus importante par rapport aux autres mouvements orofaciaux.
- **Voyelle** : Des activations additionnelles pour la tâche de vocalisation de voyelle incluent de manière bilatérale des régions du cortex auditif (cortex auditif primaire au sein du gyrus temporal transverse et partie postérieure du gyrus temporal supérieur) ainsi que le thalamus et le cortex cingulaire antérieur. Des activations dans l'hémisphère droit sont également reportées dans le cortex insulaire, les ganglions de la base (incluant le globus pallidus et le striatum dorsal), le claustrum, l'amygdale et le cortex strié.

Effet principal de la tâche motrice

L'analyse de l'effet principal de la tâche motrice (voir Figure 1.7) permet de déterminer les régions cérébrales pour lesquelles une différence significative du signal BOLD est observée pour au moins une des conditions motrices par rapport aux autres.

Des différences d'activation significatives ont ainsi été trouvées dans les gyri précentral ventral et postcentral de l'hémisphère gauche (avec deux pics d'activation localisés dans le cortex moteur primaire ventro-latéral, BA4, x/y/z : -60/-4/18, et le cortex somatosensoriel primaire, BA3, x/y/z : -

52/-8/32) et, de manière bilatérale, dans le cortex auditif (parties antérieure et postérieure du gyrus temporal transverse, BA41/42, x/y/z : -52/-16/4 et 44/-22/4).

Toutes ces régions montrent un degré d'activation plus élevé pour la tâche de vocalisation que pour les autres tâches. Comme mentionné précédemment, à cause de la séparation des tâches motrices en plusieurs runs, ces résultats nécessitent d'être interprétés avec précaution. Néanmoins, l'activation observée au niveau du cortex auditif pour la tâche de vocalisation par rapport aux autres mouvements orofaciaux reste évidemment attendue.

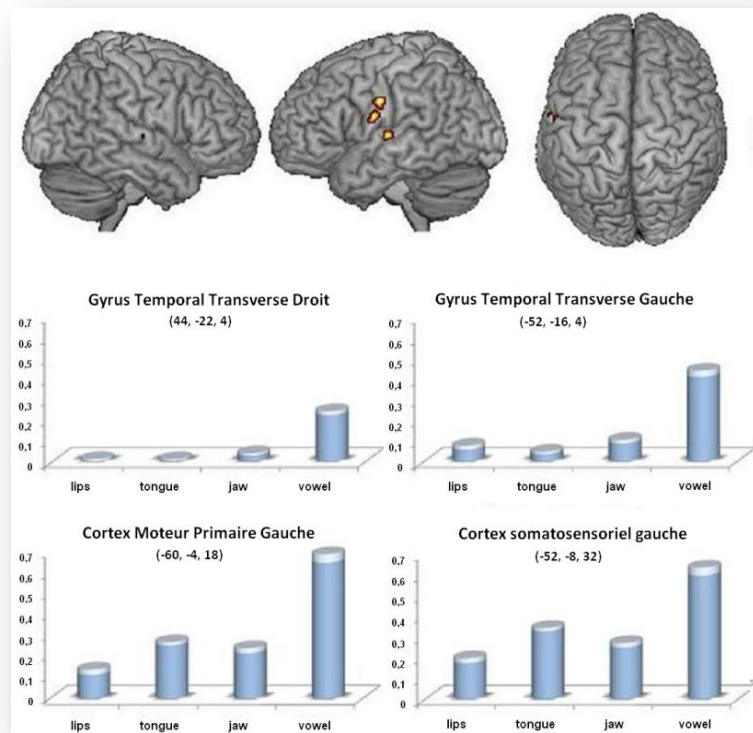


Figure 1.7 : Effet principal de la tâche motrice. Haut : vue surfacique des régions cérébrales activées de manière différenciée pour au moins une des tâches motrices par comparaison aux autres. Bas : contrastes β , reflétant le pourcentage de changement du signal BOLD par rapport à la condition de repos pour les régions activées et pour les 4 tâches motrices ($p < .05$, FWE, et taille minimale des clusters de 10 voxels).

Somatotopie sensorimotrice

Les analyses par ROI ont été réalisées individuellement à l'aide de la toolbox SPM Anatomy pour chaque participant et chaque tâche, en appliquant une correction de type "small volume correction" pour chaque ROI selon un seuil statistique défini à $p < .0001$ non corrigé. Les coordonnées MNI du pic d'activation maximale et du centre de gravité (COG) au sein des cortex primaire moteur et somatosensoriel (BA4, BA3 et BA1) ont été déterminées dans les axes médio-latéral (X), rostro-caudal (Y) et dorso-ventral (Z) et ce pour chaque participant individuellement, dans les deux hémisphères et pour toutes les conditions motrices. Les données de deux sujets ont été écartées de cette analyse, car le seuil choisi n'a pas permis de faire apparaître des activations sensorimotrices pour toutes les tâches et les deux hémisphères. Pour chaque ROI, des ANOVA à mesures répétées à deux facteurs ont été calculées. Les facteurs intra-sujet impliquaient l'hémisphère (2 niveaux : gauche, droite) et l'articulateur (4 niveaux : lèvres, langue, mandibule, larynx). Les analyses ont été effectuées pour les dimensions x, y et z indépendamment (en valeurs absolues des coordonnées MNI pour l'axe x). Pour ces analyses le niveau de significativité a été défini à $p < .05$ et une correction Greenhouse-Geisser a été appliquée en cas de non sphéricité des données. Lorsque nécessaire, des analyses post-hoc ont été effectuées en utilisant des tests de type LSD Fischer.

Pour toutes les ROIs sensorimotrices (aires 4a/p, 3a/b, 1; Geyer et al., 1996, 1999), les analyses montrent de manière quasi-systématique un effet significatif de l'Hémisphère, avec une position plus médiale, postérieure et dorsale pour l'hémisphère gauche par rapport à l'hémisphère droit. En ce qui concerne l'organisation somatotopique, une position significativement plus postérieure a été trouvée dans la dimension rostro-caudale (y) pour les lèvres par rapport à la langue, ainsi que dans la dimension dorso-ventrale (z) une position plus dorsale des lèvres que la langue et la voyelle et de la mandibule que la langue. Enfin aucune interaction entre hémisphère et articulateur n'a été observée.

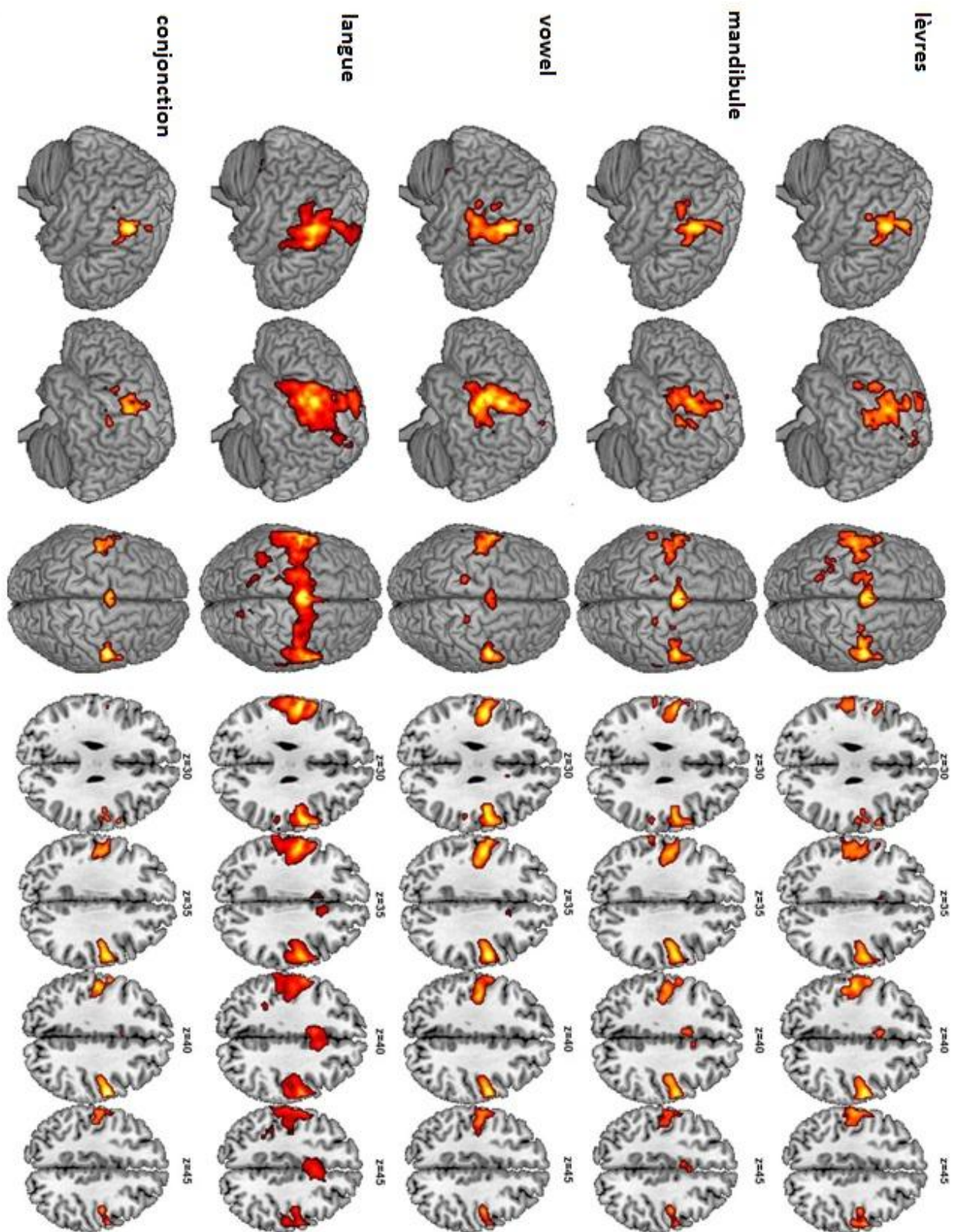


Figure 1.8 : Réseaux cérébraux des mouvements labiaux, mandibulaires, lingaux et vocaliques et réseau minimal du contrôle moteur orofacial tel que révélé par l'analyse de conjonction. Résultats reportés à un seuil corrigé (FWE, $p < .05$) et une taille minimale des clusters de 10 voxels.

Région	H	Langue					Voyelle					Mandibule					Lèvres					Conjonction				
		BA	x	y	z	T	BA	x	y	z	T	BA	x	y	z	T	BA	x	y	z	T	BA	x	y	z	T
Cortex sensorimoteur primaire	L	4	-20	-28	58	7.48	4	-60	-4	18	15.49	4	-48	-12	34	10.53	4	-48	-12	36	8.57	4	-48	-12	36	8.57
	R	4	56	-4	30	19.73	4	52	-6	32	14.67	4	54	-6	32	11.59						4	40	-12	46	7.95
	L	1,2,3	-52	-8	32	21.95	1,2,3	-52	-8	32	15.70	1,2,3	-52	-8	32	11.33	1,2,3	-54	-16	42	9.97	1,2,3	-50	-16	40	9.31
	R						1,2,3	44	-8	30	13.31	1,2,3	60	-10	34	8.30										
Régions frontales																										
Aire motrice supplémentaire	L	6	0	-6	58	20.58	6	-4	-4	62	8.93						6	-10	-2	70	6.60	6	-4	-4	62	8.93
	R						6	8	-4	60	8.18	6	2	-6	58	13.78	6	2	-6	58	11.81					
Cortex prémoteur	L	6	-62	0	26	15.50	6	-60	-4	36	11.21	6	-44	-12	60	8.97	6	-44	-12	60	9.04	6	-56	-4	40	7.56
	R						6	54	-6	38	15.55	6	54	-6	38	11.99	6	54	-6	40	12.23	6	54	-6	40	11.98
Gyrus frontal inférieur / Gyrus préfrontal	L	44/9	-58	8	30	14.76											44/9	-58	8	30	8.22					
	R	44/9	58	8	28	16.33						44	58	8	28	8.07	44/9	58	8	30	8.20	44/9	58	6	30	7.36
Régions temporales																										
Gyrus temporal transverse	L						41/42	-42	-24	8	13.43	41/42	-40	-26	12	6.71										
	R						41/42	50	-16	6	12.30															
Gyrus temporal supérieur	R					22	52	-20	0	10.71																
Régions pariétales																										
Gyrus supramarginal	L	40	-38	-44	54	9.24						40	-62	-30	34	7.52	40	-48	-26	38	9.17					
	R						40	58	-26	24	8.18	40	64	-26	22	8.50										
Opércule pariétal / Gyrus supramarginal supérieur	L	43	-62	-6	18	16.84	43	-58	-6	10	13.98	43/40	-46	-28	20	10.24	43/40	-48	-28	20	8.80	43/40	-48	-28	20	8.80
	R						43	60	-4	6	10.36	43/40	52	-28	20	7.31	43	60	-16	22	7.84					
	L	7	-20	-52	64	7.50											7	-20	-52	64	6.73					
	R	7	14	-64	56	7.84																				
Régions sous-corticales																										
Insula / opércule pariétal	R	13	36	16	8	7.01	13	46	8	2	7.06	43/13	-38	-12	18	6.96	43/13	-38	-14	20	6.86					
	L						32	10	14	34	7.14	24	-6	8	38	7.72	24	-6	8	38	8.23					
Cortex cingulaire	R						26	-2	-8	8.37																
Amgdale	L		-26	-2	-4	13.69		-26	-8	-6	10.60							-26	-2	-4	7.28		-26	-2	-4	7.28
	R		24	8	6	13.56		24	-4	-2	6.84		26	6	-6	8.00										
Pallidum	L																	-24	0	-4	8.69					
	R		18	-8	-4	9.04		22	-10	-4	8.05															
Novau caudé	R		18	4	20	7.41																				
Claustrum	R							32	-8	12	7.81															
	L							-14	-20	0	13.37							-14	-20	2	9.35					
Thalamus	R		16	-20	-2	8.21		14	-18	4	9.51		16	-20	-2	6.97										
Cervelet																										
Declive	L		-16	-62	-20	14.76		-12	-64	-16	12.95		-16	-62	-18	8.56		-16	-62	-18	6.93		-16	-62	-18	6.93
	R		18	-60	-20	13.75		20	-60	-20	12.28		20	-60	-20	7.96		18	-58	-20	7.13		18	-58	-20	7.13
Culmen	R		10	-58	-14	8.66																				

Table 1.1 : Coordonnées dans l'espace MNI des pics d'activation maximale observés pour les quatre tâches motrices orofaciales et communs à ces 4 tâches (analyse de conjonction). Résultats reportés à un seuil corrigé (FWE, $p < .05$) et une taille minimale des clusters de 10 voxels.

ROIs sensorimotrices	Hémisphère gauche												Hémisphère droit											
	lèvres			mandibule			voyelle			langue			lèvres			mandibule			voyelle			langue		
	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
Pics d'activation																								
BA 4a,p	-46	-12	37	-43	-14	36	-43	-13	37	-47	-11	35	51	-8	39	48	-8	37	51	-6	35	51	-5	34
BA 3a,b	-46	-16	38	-46	-15	36	-50	-10	31	-50	-12	33	48	-11	37	51	-9	33	55	-6	28	56	-5	28
BA 1	-52	-14	43	-52	-12	41	-53	-11	39	-56	-10	36	54	-10	41	57	-7	36	59	-5	31	60	-4	31
Coordonnées moyennes	-48	-14	40	-47	-13	38	-49	-11	36	-51	-11	34	51	-10	39	52	-8	35	55	-6	32	55	-5	31
Centres de gravité (COG)																								
BA 4a,p	-44	-11	39	-43	-12	39	-42	-12	39	-45	-10	36	47	-8	38	47	-8	37	49	-7	36	49	-6	34
BA 3a,b	-45	-16	39	-47	-14	36	-50	-10	31	-49	-11	33	48	-12	36	51	-9	33	55	-6	28	54	-7	28
BA 1	-54	-13	44	-54	-11	42	-55	-11	40	-57	-9	38	55	-11	41	58	-7	35	60	-6	31	60	-7	33
Coordonnées moyennes	-48	-14	40	-48	-12	39	-49	-11	37	-50	-10	36	50	-10	38	52	-8	35	55	-6	32	54	-7	32
Références																								
Corfield et al. (1999) * - mouvement lingual horizontal										-63	-7	24										64	-5	19
Lotze et al. (2000a) * - protrusion labiale, mouvement lingual vertical	-53	-18	40							-53	0	35	55	-8	41							67	-3	26
Lotze et al. (2000b) * - protrusion labiale	-59	-16	43										53	-10	39									
Riecker et al. (2000) - mouvement lingual horizontal										-63	-3	30										57	-3	24
Stippich et al. (2002) * - mouvement lingual vertical										-54	-12	34										53	-8	34
Fesl et al. (2003) * - mouvement lingual horizontal										-63	-8	37										67	-8	32
Gerardin et al. (2003) * - protrusion labiale	-58	-15	45										61	-9	52									
Shinagawa et al. (2003) * - protrusion linguale										-48	-14	34										51	-8	37
Shinagawa et al. (2003) * - mouvement lingual droite										-48	-14	36										51	-8	37
Shinagawa et al. (2003) * - mouvement lingual gauche										-48	-14	38										59	-1	24
Hesselmann et al. (2004) * - protrusion labiale, mouvement lingual horizontal	-41	-21	44							-55	-12	27	53	-21	49							58	-13	30
Hanakawa et al. (2005) * - étirement labial unilatéral	-54	-8	41										57	-3	40									
Pulvermüller et al. (2006) –mouvement lingual ou labial haut/bas	-48	-10	36							-56	-8	28												
Terumitsu et al. (2006) – mouvement lingual ou phonation de voyelle							-40	-19	42	-55	-8	32						51	-11	38		59	-6	30
							-56	-3	21															
Vincent et al. (2006) * - mouvement lingual horizontal/vertical																						61	-1	26
Brown et al. (2008) * – protrusion labiale et mouvement lingual vertical	-52	-14	36							-64	-12	30	58	-12	34							58	-14	34
Brown et al. (2008) * – coups de glotte							-38	-16	34									44	-12	36				
Brown et al. (2008) * – phonation de voyelle							-40	-12	32									44	-10	36				
Stippich et al. (2008) * - mouvement lingual vertical										-54	-16	41										54	-11	39
Coordonnées moyennes	-52	-15	41				-44	-13	32	-55	-10	34	56	-11	43				46	-11	37	58	-7	31

Table 1.2 : Somatotopie sensorimotrice. Valeurs moyennes des centres de gravité (COG) et des pics d'activation maximale observées dans les analyses ROI et revue des pics d'activations principaux observés dans le cortex sensorimoteur primaire lors de l'exécution de simples mouvements orofaciaux dans de précédentes études (espace MNI).

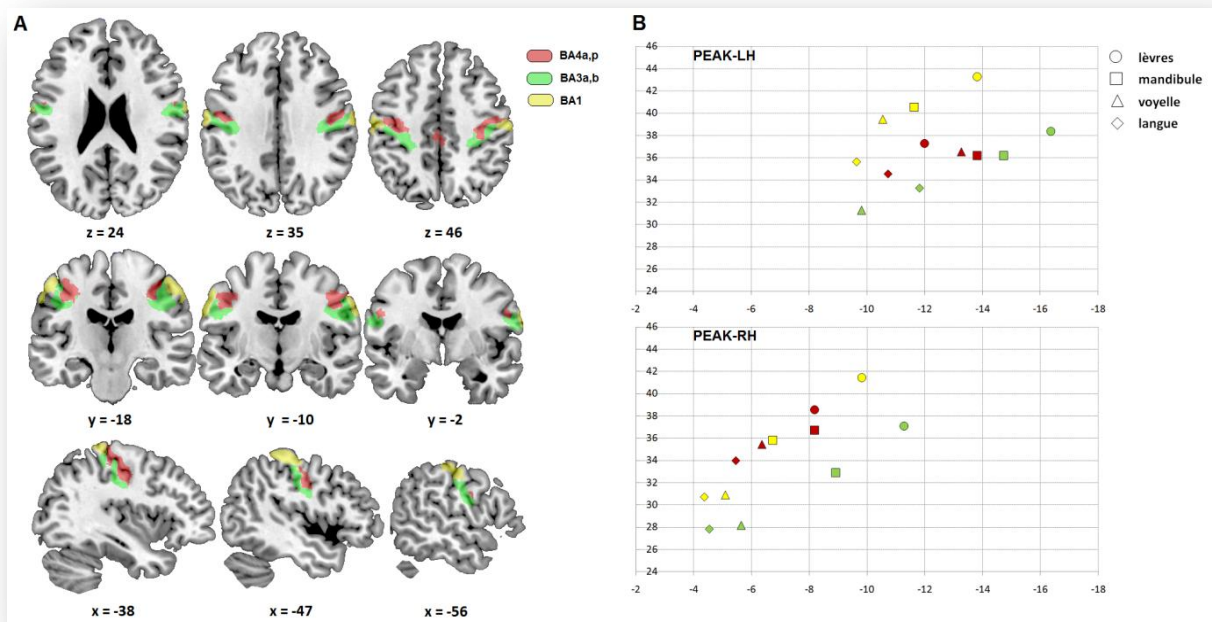


Figure 1.9 : Somatotopie sensorimotrice (analyse par régions d'intérêts). A) ROI correspondantes aux aires BA 4a/4p (rouge), BA 3a/3b (vert) et BA 1 définies via la Toolbox SPM Anatomy. B) Projections des coordonnées Y et Z des pics d'activation maximale des quatre articulateurs. Une organisation somatotopique est observée, impliquant une position plus postérieure des lèvres par rapport à la langue dans la dimension rostro-caudale (Y) et une position plus dorsale des lèvres par rapport à la langue et la vocalisation ainsi que pour la mandibule par rapport à la langue dans l'axe dorso-ventral (Z).

Dimension médio-laterale (x)

- **BA4** : Des effets significatifs ont été observés pour le facteur Hémisphère et Articulateur ($F_{(1,10)} = 24.7$; $p < .001$; $F_{(3,30)} = 5.1$; $p < .006$), avec une position plus médiale dans l'hémisphère gauche que dans l'hémisphère droit (45 vs 50 mm) et une position plus latérale pour les lèvres et la langue que pour la voyelle et la mandibule (48, 49, 46, 47mm, respectivement).
- **BA3** : Des effets significatifs des facteurs Hémisphère et Articulateur sont observés ($F_{(1,10)} = 20.1$; $p < .001$; $F_{(3,30)} = 8.7$; $p < .001$), avec une position plus médiale dans l'hémisphère gauche (48 vs 53mm) et une position plus latérale pour la voyelle et la langue par rapport aux lèvres et à la mandibule (53, 53, 48, 48mm, respectivement).
- **BA1** : Des effets significatifs des facteurs Hémisphère et Articulateur sont observés ($F_{(1,10)} = 21.9$; $p < .001$; $F_{(3,30)} = 3.8$; $p < .02$), avec une position plus médiale dans l'hémisphère gauche (53 vs 57mm) et une position plus latérale pour la langue par rapport aux lèvres et à la mandibule (58, 53, 54mm, respectivement).

Dimension rostro-caudale (y)

- **BA4** : Des effets significatifs des facteurs Hémisphère et Articulateur sont observés ($F_{(1,10)} = 31.4$; $p < .001$; $F_{(3,30)} = 6.6$; $p < .001$), avec une position plus postérieure dans l'hémisphère gauche par rapport à l'hémisphère droit (-12 vs -7mm) et une position plus antérieure pour la langue par rapport aux lèvres, la mandibule et la voyelle (-8, -10, -11, -10mm, respectivement).
- **BA3** : Des effets significatifs des facteurs Hémisphère et Articulateur sont observés ($F_{(1,10)} = 45.9$; $p < .001$; $F_{(3,30)} = 12.9$; $p < .001$) avec une position plus postérieure dans l'hémisphère gauche (-13 vs -8mm) et une position plus antérieure pour la voyelle et la langue par rapport aux lèvres et la mandibule (-8, -8, -14, -12mm, respectivement).
- **BA1** : Des effets significatifs des facteurs Hémisphère et Articulateur sont observés ($F_{(1,10)} = 19.4$; $p < .001$; $F_{(3,30)} = 4.7$; $p < .008$), avec une position plus postérieure dans l'hémisphère gauche (-11

vs -6mm) et une position plus antérieure pour la langue et la voyelle par rapport aux lèvres (-8, -7, -12mm, respectivement).

Dimension dorso-ventrale (z)

- **BA4** : Un effet significatif du facteur Articulateur est observé ($F_{(3,30)} = 7.0$; $p < .001$), avec une position plus dorsale pour les lèvres, la mandibule et la voyelle par rapport à la langue (38, 36, 36, 34mm, respectivement) tout comme une position plus dorsale pour les lèvres par rapport à la voyelle. L'effet de l'hémisphère n'est pas significatif.
- **BA3** : Des effets significatifs des facteurs Hémisphère et Articulateur ont été observés ($F_{(1,10)} = 5.3$; $p < .05$; y : $F_{(3,30)} = 11.1$; $p < .001$), avec une position plus dorsale dans l'hémisphère gauche (35 vs 31mm) et une position plus dorsale pour les lèvres et la mandibule par rapport à la voyelle et la langue (38, 35, 30, 31mm, respectivement). L'interaction n'est pas significative.
- **BA1** : Des effets significatifs des facteurs Hémisphère et Articulateur ont été observés ($F_{(1,10)} = 10.2$; $p < .009$; y : $F_{(3,30)} = 7.5$; $p < .001$), avec une position plus dorsale pour les lèvres par rapport à la voyelle et la langue (42, 35, 33mm, respectivement) et une position plus dorsale pour la mandibule par rapport à la langue (35, 33mm, respectivement).

Activités cérébrales liées aux délais d'acquisition

Régions	Effet principal					Contrastes estimés		
	x	y	z	F		4s	5s	6s
Régions frontales								
Aire motrice supplémentaire	6	-8	6	52	15.46	0.11	0.08	0.02
	6	8	-4	62	20.98	0.14	0.14	0.07
Cortex prémoteur	6	64	2	10	14.25	0.02	0.12	0.13
Cortex moteur primaire	4	-36	-14	48	25.74	0.11	0.11	0.05
	4	44	-12	42	20.64	0.14	0.19	0.12
Gyrus frontal moyen	8	0	18	50	9.95	0.03	0.09	0.02
Régions temporales								
Gyrus temporal supérieur	22	-52	-2	4	13.38	0.05	0.16	0.14
Régions pariétales								
Cortex somatosensoriel	3	-38	-18	36	16.78	0.20	0.22	0.15
	3	24	-26	58	16.89	0.6	0.10	0.04
Opércule pariétal	43	-48	-6	10	26.04	0.04	0.11	0.11
Régions sous-corticales								
Insula	13	-38	2	14	11.74	0.06	0.11	0.09
Amygdale	22	-10	-8	13.33	-0.02	0.04	0.03	
Hippocampe	22	2	-18	17.30	-0.03	0.07	0.07	
Cervelet								
Declive	-14	-66	-10	39.88	-0.03	0.09	0.08	
	28	-70	-22	36.34	-0.03	0.09	0.10	
Culmen	2	-50	0	29.04	-0.10	0.07	0.06	
Régions occipitales								
Gyrus calcarine	-18	-66	10	32.30	0.01	0.09	0.02	
	4	-78	10	23.64	-0.02	0.08	0.01	
Gyrus lingual	16	-66	-10	26.69	-0.08	0.06	0.02	

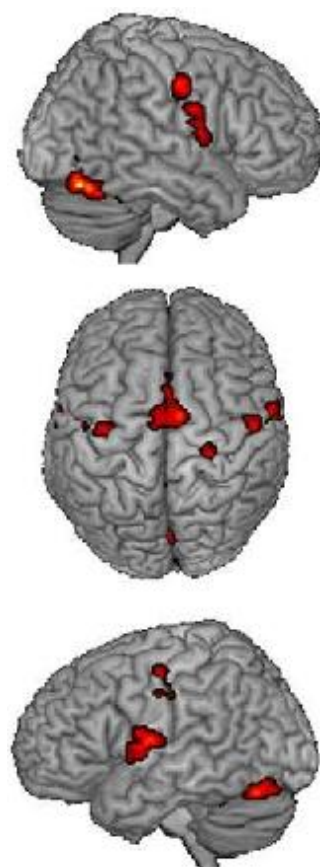


Table 1.3 et Figure 1.10 : Effet principal du délai d'acquisition. Principales régions cérébrales activées de manière différenciée pour au moins un des 3 délais d'acquisition. Les estimations de contrastes indiquent les variations du signal par rapport à la condition de repos pour les 3 délais ($p < .05$, FWE, et taille minimale des clusters de 10 voxels).

Chaque délai d'acquisition des images fonctionnelles a été individuellement contrasté par rapport à la ligne de base. L'analyse de groupe révèle des activations significatives largement communes aux 3 délais d'acquisitions (4s, 5s, 6s), confirmant par là même la bonne adéquation de ces délais dans l'acquisition du signal BOLD lors des tâches motrices. Néanmoins, l'effet principal des trois délais

d'acquisition montre également des décours temporels significativement différents dans certaines régions corticales et sous-corticales (voir Table 1.3 et Figure 1.10).

- Des courbes en cloche de l'activité neuronale à travers les trois délais d'acquisition sont observées (pic d'activation maximal à 5s et significativement supérieur à ceux observés à 4s et 6s) dans le cortex sensorimoteur primaire ventral bilatéral, l'insula gauche et le gyrus temporal supérieur gauche, ainsi que l'amygdale droite et le gyrus calcarine.
- Une décroissance de l'activation neuronale de 4s à 6s est observée au sein de l'aire motrice supplémentaire, s'étendant au gyrus frontal supérieur.
- Une augmentation de l'activation neuronale de 4s à 6s est observée dans l'opercule pariétal gauche, au niveau du cortex prémoteur ventral, de l'hippocampe et du gyrus parahippocampique droits, ainsi que dans le vermis du cervelet (déclive et culmen).

1.1.4 Discussion

L'objectif de la présente étude était de clarifier les réseaux neuronaux sous-jacents au contrôle moteur de mouvements simples supralaryngés et laryngés/vocaliques et d'affiner l'organisation somatotopique des représentations motrices et somatosensorielles labiales, mandibulaires, vocaliques/laryngées et linguales. Pour ce faire, une méthode d'acquisition des images fonctionnelles en sparse sampling a été utilisée dans le but de minimiser de possibles artefacts du champ magnétique liés aux mouvements orofaciaux.

Trois résultats majeurs ont été observés.

- Les mouvements orofaciaux ont activé un ensemble de structures corticales et sous-corticales communes formant un réseau neural minimal impliqué dans le contrôle moteur orofacial.
- En dehors des activations temporelles auditives pour la vocalisation de la voyelle /i/, les différences entre tâches motrices ont été restreintes aux régions ventro-latérales du cortex moteur et somatosensoriel primaire de l'hémisphère gauche, avec une intensité de signal plus ample pour la tâche de vocalisation. Ce résultat reflète vraisemblablement un degré de coordination neuromusculaire et de traitement du feedback sensoriel plus complexe par rapport aux autres tâches.
- En utilisant des analyses individuelles par régions d'intérêt, une organisation somatotopique séquentielle dorso-ventrale a été observée au sein des aires primaires motrices et somatosensorielles pour les mouvements labiaux, mandibulaires, vocaliques/laryngés et linguaux respectivement.

Avant de discuter de ces résultats, il est néanmoins important de souligner une limite importante de la présente étude. Comme noté précédemment, bien que la tâche de vocalisation ait été choisie *a posteriori* afin de déterminer l'activité motrice laryngée, elle requiert également des activités musculaires linguale et labiale. De même, de par la connexion anatomique entre la mandibule et la langue, les mouvements mandibulaires peuvent induire des mouvements linguaux (même si ceux-ci sont passifs). De là, des contributions mutuelles entre articulateurs orofaciaux lors de la vocalisation de la voyelle /i/ ne peuvent bien entendu pas être entièrement exclues.

Réseau orofacial minimal

Le réseau neural commun aux mouvements supralaryngés et vocaliques révèle des fortes activations bilatérales au sein du cortex sensorimoteur (en incluant les cortex primaires moteur et somatosensoriel et les parties ventrales et dorsales du cortex prémoteur), l'aire motrice supplémentaire et les hémisphères cérébelleux supérieurs. Des activations ont également été

trouvées dans le gyrus frontal inférieur postérieur (partie operculaire) droit, l'opercule pariétal gauche et une région du lobule pariétal inférieur adjacent et, enfin, le striatum dorsal gauche. Ce réseau neural minimal des actions motrices orofaciales est tout à fait concordant avec de précédentes études IRMf portant sur le contrôle moteur orofacial. Notamment, une activité neurale commune pour la mandibule et la langue (Dhanjal et al., 2008) a été observée au niveau du cortex prémoteur ventral latéral, de l'opercule frontal ventral, de l'insula, du cortex primaire moteur et somatosensoriel, de l'aire motrice supplémentaire et du cervelet (lobule VI). Lotze et collaborateurs (2000a) ont rapporté des résultats similaires pour les mouvements linguaux et labiaux bien que, de manière surprenante, aucune activation n'ait été observée au niveau de l'aire motrice supplémentaire.

Par rapport aux actions orofaciales complexes coordonnées, nos résultats sont également cohérents avec les études portant sur la déglutition et la respiration (Sawczuk et Mosier, 2001; Loucks et al., 2007), de vocalisation (Jürgens 2002, 2009) ainsi que sur la production ouverte de la parole (Riecker et al., 2000, 2005, 2008; Brown et al., 2005; Bohland et Guenther, 2006; Guenther, Ghosh et Tourville 2006; Sörös et al., 2006). Premièrement, il a été montré que le contrôle neural de la déglutition et de la respiration dépendent de structures neurales communes, incluant le cortex bilatéral sensorimoteur primaire, l'aire motrice supplémentaire, le thalamus, le cervelet, le noyau caudé, le globus pallidus et le bulbe rachidien (Sawczuk et Mosier, 2001; voir aussi McKay et al., 2003). En accord avec nos résultats, de nombreuses études ont mené à suggérer un réseau minimal de la production de parole ouverte (Bohland et Guenther, 2006), qui comprend les structures mésio-frontales (aire motrice supplémentaire et gyrus cingulaire antérieur), les convolutions pré- et postcentrales bilatérales, s'étendant rostralement aux parties postérieures du gyrus frontal inférieur, l'insula antérieure gauche tout comme les structures des ganglions de la base bilatéralement (notamment le putamen et le globus pallidus), le cervelet (notamment le lobule VI, incluant le déclive), le thalamus et le gyrus temporal supérieur (Brown et al., 2005; Bohland et Guenther, 2006; Guenther, Ghosh et Tourville 2006; Sörös et al., 2006; Riecker et al., 2005). Enfin, Chang et collaborateurs (2009) ont montré des activations communes lors de mouvements du conduit vocal liés et non liés à la parole au niveau du gyrus frontal inférieur, du cortex prémoteur ventral, de l'aire motrice supplémentaire, du gyrus temporal supérieur, de l'insula, du gyrus supramarginal, du cervelet et des ganglions de la base. Finalement, le réseau minimal orofacial observé apparaît également en accord avec les substrats neuronaux du chant (voir Brown et al., 2005 pour revue) et du sifflement (Dresel et al., 2005), malgré des différences d'activation au niveau de l'amygdale (non rapporté pour le chant) et de l'insula (non rapporté pour le sifflement).

En regard de ces études, nos résultats permettent ainsi d'affiner le réseau neuronal minimal impliqué dans le contrôle moteur orofacial spécifique ou non à la production de la parole. Ce réseau englobe le cortex ventral sensorimoteur et le cortex prémoteur ventral/dorsal de manière bilatérale, le gyrus frontal inférieur postérieur (partie operculaire) droit, l'aire motrice supplémentaire, l'opercule pariétal gauche et le lobule pariétal inférieur adjacent, les ganglions de la base et le cervelet. Comme souligné précédemment, les fonctions de ces régions corticales et sous-corticales peuvent être globalement assignées à des boucles de préparation, de planification, d'exécution et de coordination motrice (pour une revue, voir Jürgens, 2002; Riecker et al., 2005).

Effet principal de la tâche motrice

Dans la présente étude, les réseaux neuronaux impliqués dans les mouvements orofaciaux simples sont largement partagés et incluent des régions corticales et sous-corticales similaires. En dehors des gyri temporaux supérieurs bilatéraux pour la tâche de vocalisation, la seule différence significative entre les quatre tâches motrices a été observée au niveau du cortex sensorimoteur ventro-latéral, avec des activations plus amples pour la vocalisation par rapport aux autres tâches.

De par la séparation des tâches supralaryngées et laryngée/vocalique en runs différents, les activations sensorimotrices accrues lors de la tâche de vocalisation doivent cependant être interprétées avec précaution. Il est bien sûr tentant de supposer que ces régions soient reliées spécifiquement aux mouvements laryngés lors de la production de la parole. Dans ce cas, les pics d'activation observés dans les analyses par ROI lors de la vocalisation et du mouvement lingual (voir ci-dessous, section "Somatotopie articulatoire"), légèrement plus dorsaux que les activations présentes, pourraient refléter la contribution d'activité musculaire linguale lors de la vocalisation, plutôt que l'activité purement laryngée. D'un autre côté, ces deux pics d'activations dans le cortex primaire ventro-latéral moteur et somatosensoriel ont également été activés lors des mouvements mandibulaires et linguaux, mais dans une moindre mesure/ampleur. Bien que les mouvements de rétraction linguale induisent aussi un feedback sensoriel important, une interprétation alternative serait que cette activité sensorimotrice primaire accrue pourrait être due à des interactions sensorimotrices plus complexes requises pour la vocalisation maintenue de la voyelle. Afin d'atteindre et maintenir une cible de son de parole précise, ces activations accrues des cortex ventrolatéral primaire gauche moteur et somatosensoriel pourraient en ce cas refléter une implication plus importante des mécanismes de contrôle moteur en ligne et de traitement du feedback somatosensoriel et auditif (Guenther, 2006; Bohland et Guenther, 2006; Guenther et Vladusich, 2012).

Somatotopie articulatoire

Du fait d'activations sensorimotrices primaires communes aux quatre articulateurs, aucune organisation somatotopique n'a pu être révélée par l'analyse de groupe. Pour mieux prendre en compte les différences individuelles anatomiques et fonctionnelles entre participants, des analyses par régions d'intérêts ont été effectuées pour évaluer l'organisation somatotopique de manière individuelle pour les articulateurs laryngé et supralaryngés. Dans l'ensemble, les coordonnées moyennes observées pour les représentations labiales, mandibulaires, linguales et vocaliques dans les deux hémisphères et les trois régions d'intérêt sensorimotrices correspondent avec celles rapportées par des études IRMf précédentes sur des gestes articulatoires simples. Pour toutes les régions d'intérêt sensorimotrices, une position plus postérieure émerge dans la dimension rostro-caudale (y) pour les lèvres par rapport à la langue et une position plus dorsale dans la dimension dorso-ventrale (z) pour les lèvres par rapport à la langue et la voyelle, de même que pour la mandibule par rapport à la langue. Le résultat principal de ces analyses par ROI est ainsi celui d'une organisation somatotopique dorso-ventrale des lèvres, de la mandibule, de la voyelle et de la langue, avec les représentations des mouvements de lèvres occupant la position la plus dorso-caudale, celles des mouvements vocaliques/laryngés et linguaux les positions les plus ventro-rostrales et celles de la mandibule positionnées dorsalement par rapport à la langue.

- Ces résultats apparaissent cohérents avec la position plus dorsale des lèvres par rapport à la langue dans les homunculi moteur et somatosensoriel, comme démontré par de précédentes études de stimulation électrocorticale et d'IRMf (Penfield et Rasmussen, 1950; Lotze et al., 2000a; Hesselmann et al. 2004; Pulvermüller et al. 2006).
- En accord avec la cartographie électrocorticale élaborée par Penfield et Rasmussen (1950), nous avons également observé l'aire du larynx positionnée plus ventralement que celle des lèvres et plus dorsalement que l'aire du larynx et de la langue (bien qu'aucune différence significative n'ait été observée entre la position des lèvres par rapport à la mandibule et que la seule différence significative entre les positions de la mandibule et de la voyelle ait été observée dans l'aire BA3)³⁶.

³⁶ Une précédente étude IRMf a également montré une position légèrement plus médiale et dorsale pour la mâchoire (mouvements d'ouverture/fermeture) par rapport à la langue (mouvement vertical) dans le cortex sensorimoteur primaire, malheureusement sans faire part des coordonnées précises des pics d'activations séparément pour ces deux articulateurs (Dhanjal et al., 2008).

- Enfin, la localisation motrice des mouvements vocaliques/laryngés s'avère plus dorsale que la position de la langue. Brown et collaborateurs (2008) rapportent des pics d'activation similaires dans le cortex moteur gauche pour les mouvements laryngés vocaux et non-vocaux (abduction de cordes vocales et vocalisation de la voyelle schwa) en les localisant entre les pics d'activation observés pour les mouvements des lèvres et de la langue. Néanmoins, ils rapportent une position plus dorsale pour l'aire du larynx par rapport aux lèvres et aussi à la langue dans l'hémisphère droit. Nous observons que, dans leur étude, les pics d'activation des lèvres et de la langue dans le cortex moteur primaire droit sont quasiment identiques, un résultat qui ne correspond pas aux études précédentes qui montrent l'aire des lèvres positionnée en-dessus de l'aire de la langue dans les deux hémisphères (Lotze et al., 2000a; Hesselmann et al. 2004; voir Table 1.2). Il est également à noter une étude montrant dans le cortex moteur primaire de l'hémisphère gauche lors d'une tâche de production de voyelle deux pics d'activation spécifiques et localisés de manière plus dorsale et plus ventrale par rapport à un pic d'activation lié à un mouvement lingual (Terumitsu et al., 2006).

Enfin, quel que soit l'articulateur, les résultats montrent une asymétrie hémisphérique significative des représentations sensorimotrices, avec des sites d'activation positionnés de manière plus médiale, postérieure et dorsale dans l'hémisphère gauche par rapport à l'hémisphère droit. Une revue des études IRMf portant sur les mouvements simples des lèvres, de la langue, et du larynx montre une asymétrie hémisphérique sensorimotrice similaire pour tous les articulateurs, avec une position plus médiale et plus postérieure des sites d'activation dans l'hémisphère gauche. Ce résultat pourrait donc suggérer des asymétries anatomiques et/ou fonctionnelles au sein du cortex sensorimoteur primaire (Sowman et al., 2009) sans distinction de la prévalence manuelle ou du sexe des participants (Amunts et al., 2000).

Patterns temporels d'activations

En accord avec de précédentes études utilisant la méthode de sparse sampling pour l'étude de tâches motrices orofaciales et de production de parole (Dresel et al., 2005; Riecker et al., 2005; Gracco, Tremblay et Pike, 2005; Bohland et Guenther, 2006; Özdemir, Norton et Schlaug, 2006; Sörös et al., 2006), les données fonctionnelles ont été acquises environ 4s, 5s ou 6s après le début des mouvements orofaciaux afin de capturer le pic maximal des réponses hémodynamiques associées (Dresel et al., 2005; Riecker et al., 2005; Özdemir, Norton et Schlaug, 2006). Néanmoins, des différences entre les patterns temporels d'activation correspondant aux mouvements orofaciaux (sans distinctions de tâches) ont été observées. Ainsi, des réponses hémodynamiques avec un pic maximal à 4s ont été observées au niveau de l'aire motrice supplémentaire et du gyrus frontal supérieur. Les régions cérébrales montrant une augmentation d'activité hémodynamique de 4s à 6s impliquaient le cervelet, l'opercule pariétal gauche, le cortex prémoteur, l'hippocampe et le gyrus parahippocampique droit. Toutes les autres régions cérébrales ont montré un décours temporel "en cloche" (avec un pic d'activation maximal à 5s) ou aucune différence. En somme, ces résultats confirment l'adéquation de ces trois délais d'acquisition pour une capture optimale des pics maximaux des réponses hémodynamiques des régions cérébrales impliquées dans le contrôle moteur orofacial. Ces résultats apparaissent également cohérents avec des études IRMf récentes utilisant la méthode de sparse sampling lors de la répétition de syllabes (Riecker et al., 2005), de tâches de sifflement et de respiration (Dresel et al., 2005). Ces études montrent en effet des réponses hémodynamiques séquentielles à travers les régions cérébrales, avec notamment des activations plus précoces au niveau de l'aire motrice supplémentaire et des activations plus tardives au niveau du cervelet (bien que Riecker et collaborateurs (2005) observent des différences inverses entre activations précoces/tardives dans le cervelet droit supérieur/inférieur). Les auteurs de ces études émettent de ces résultats la possibilité que ces différences temporelles correspondent à l'ordre d'activation des régions cérébrales dans l'initiation et dans l'exécution de l'acte moteur. Dans la présente étude, il ne s'agit pas d'émettre des hypothèses semblables, étant donné que les

différences des points temporels d'acquisition varient de 1s ou 2s, alors que la durée des processus d'initiation et d'exécution motrice est au maximum de l'ordre de quelques dizaines de millisecondes. De plus, ce type d'inférence doit être considéré avec précaution, étant donné que les relations entre décharge neuronale et réponse hémodynamique (en termes de latence, amplitude et de durée) ne sont pas linéaires, ni clarifiées. Les événements hémodynamiques correspondent en effet à des échelles de temps de l'ordre de quelques secondes, alors que les événements neuronaux opèrent au niveau de la milliseconde. De plus, nous l'avons vu, les paramètres du signal BOLD dépendent non seulement de l'activité neuronale, mais aussi de la distance entre la région activée et l'artère qui l'irrigue, tout comme le diamètre de celle-ci (Logothetis et al. 2001). Néanmoins, ces résultats fournissent des informations importantes sur le déroulement temporel des réponses hémodynamiques des différentes structures cérébrales impliquées dans les tâches de contrôle moteur orofacial, un résultat potentiellement intéressant pour de futures études IRMf dans l'optimisation des paramètres d'acquisition et une capture optimale des réponses hémodynamiques des régions cérébrales spécifiquement associées aux mouvements orofaciaux.

1.1.5 Conclusion

Les résultats de cette étude ont permis de mettre en évidence un réseau neural fonctionnel minimal contrôlant les mouvements laryngés et supralaryngés, incluant le cortex primaire sensorimoteur et prémoteur, le gyrus frontal inférieur droit, l'aire motrice supplémentaire, l'opercule pariétal gauche et le lobule pariétal inférieur adjacent, les ganglions de la base et le cervelet. En dehors du cortex temporal auditif activé lors de la tâche de vocalisation, des différences d'activité entre tâches motrices ont été trouvées dans le cortex sensorimoteur ventrolatéral gauche, avec une intensité de signal plus ample pour la vocalisation. Enfin, bien qu'aucune somatotopie sensorimotrice n'ait pu être observée pour les quatre articulateurs lors de l'analyse de groupe, une organisation somatotopique séquentielle dorso-ventrale a néanmoins été trouvée pour les mouvements labiaux, mandibulaires, vocaliques et linguaux au sein du cortex primaire moteur et somatosensoriel en effectuant des analyses individuelles par ROI. Le réseau neural minimal associé aux mouvements orofaciaux et l'organisation séquentielle somatotopique dorso-ventrale des articulateurs observée dans la présente étude pourraient permettre une meilleure spécification des modèles (neuro)computationnels de production de la parole et apportent des résultats potentiellement importants pour de futures études de neuroimagerie portant notamment sur le contrôle moteur orofacial, la production et les troubles de la parole, tout comme la réorganisation cérébrale post-traumatique.

Chapitre 1.2

Corrélat neuronal des systèmes de perception et de production des voyelles du Français

Résumé / Abstract

Certains modèles neurobiologiques récents de la perception et de la production de la parole postulent l'existence d'un couplage fonctionnel entre systèmes sensoriels et moteur. Lors de la perception de la parole, des processus moteurs inférentiels permettraient de contraindre l'interprétation phonétique du signal acoustique perçu, notamment lors de conditions d'écoute défavorables. Lors de la production de la parole, des boucles correctives sensorimotrices permettraient, par l'examen de possibles discordances entre états sensoriels prévus, de contrôler en ligne la réalisation des actes produits. Ces boucles seraient notamment impliquées dans le cas d'un retour sensoriel altéré et/ou lors de la production de séquences articulatoires complexes. La présente étude d'imagerie par résonance magnétique fonctionnelle avait pour but de tester si la perception et la production de voyelles isolées induisent également de telles interactions sensorimotrices, indépendamment de mécanismes de séquençage syllabique ou de coarticulation et en utilisant la technique d'acquisition sparse sampling dans le but de limiter le bruit du scanner. Nos résultats montrent qu'une voie postéro-dorsale gauche, liant les percepts auditifs de la parole avec leurs représentations articulatoires et incluant le gyrus frontal inférieur, le cortex prémoteur adjacent et la jonction temporo-pariétale, est communément activée lors de la perception et de la production des voyelles. Des analyses spécifiques portant sur les traits phonétiques confirment l'implication de cette voie postéro-dorsale dans le traitement sensoriel et le contrôle moteur des voyelles mais ne permettent pas de révéler une ségrégation topographique significative entre traits. Pris ensemble, ces résultats suggèrent une organisation largement distribuée des représentations cérébrales des voyelles au sein de ces régions sensorimotrices et apportent des arguments supplémentaires en faveur d'un couplage fonctionnel entre systèmes de perception et de production de la parole.

Recent neurobiological models of speech perception and production argue for a functional coupling between sensory and motor systems. In speech perception, motor inferential processes are assumed to constrain the phonetic interpretation of the acoustic speech signal, and to be particularly involved under adverse listening conditions. In speech production, sensorimotor corrective loops are thought to help registering possible discrepancy between the intended and the actual sensory states in order to further control production, and to be strongly recruited in case of altered auditory feedback and/or complex articulatory speech sequences. In the present functional magnetic resonance imaging study, we tested whether isolated vowel perception and production might also induce sensorimotor interactions, independently of any syllable sequencing and coarticulation mechanisms and using a sparse acquisition technique in order to limit influence of scanner background noise. Our results demonstrate that a left postero-dorsal stream, linking auditory speech percepts with articulatory representations and including the posterior inferior frontal gyrus, the adjacent ventral

premotor cortex and the temporo-parietal junction, is an influential part of both vowel perception and production. Specific analyses on phonetic features further confirmed the involvement of the left postero-dorsal stream in vowel processing and motor control but did not reveal clear topographic segregation between vowel features. Altogether, these results strongly suggest that vowel representations are largely distributed over sensorimotor brain areas and provide further support for a functional coupling between speech perception and production systems.

- *Mots clés : Perception et production de la parole, interactions sensorimotrices, voyelle, traits phonétiques, voie dorsale, IRMf, sparse sampling.*
- *Keywords: speech perception and production, sensorimotor interactions, vowel, phonetic features, dorsal pathway, fMRI, sparse sampling.*

1.2.1 Introduction

La parole est construite sur un jeu de correspondances entre représentations sensorielles et motrices, notamment lors de l'acquisition du langage les premières années de vie. Cette co-structuration des représentations sensorielles et motrices des unités de parole s'établirait de manière progressive chez l'enfant au travers de mécanismes sensorimoteurs d'apprentissage orofacial et d'imitation vocale. Ces premières étapes exploratoires d'acquisition de la parole permettraient l'établissement d'associations entre buts moteurs et sensoriels, notamment auditifs ("modèle direct") et, de là, une meilleure spécification des actes moteurs en fonction des cibles auditives de son environnement langagier ("modèle inverse").

Chez l'adulte, bien que la persistance d'un tel couplage fonctionnel entre systèmes sensoriels et moteur reste débattue, de nombreux modèles phonétiques, psycholinguistiques et neurobiologiques de la perception et/ou de la production de la parole défendent également le rôle fondamental des interactions entre systèmes sensoriels et moteur lors de la perception et production de la parole (Hickok et Poeppel, 2000, 2004, 2007; Scott et Johnsrude, 2003; Callan et al., 2004; Guenther, 2006; Wilson et Iacoboni, 2006; Skipper et al., 2007; Rauschecker et Scott, 2009; Rauschecker, 2011; Hickok, Houde et Rong, 2011; Guenther et Vladusich, 2012; Perkell, 2012; Schwartz et al., 2012; pour une revue, voir la partie introductive de cette thèse). Lors de la production de parole, la modulation des réponses neuronales observées au sein des cortex auditif et sensoriels refléterait des mécanismes de contrôle des retours sensoriels, permettant la comparaison des conséquences sensorielles prédites des actes moteurs de parole et des entrées sensorielles réelles et ainsi un contrôle en ligne des productions de parole (pour des revues récentes, voir Perkell, 2012; Guenther et Vladusich, 2012). Inversement, l'implication du système moteur observée lors de la perception de la parole pourrait correspondre à des mécanismes de simulation motrice dont le but serait de contraindre l'interprétation phonétique finale des entrées sensorielles par la génération interne de possibles candidats articulatoires (pour des revues récentes, voir Schwartz, Sato et Fadiga, 2008; d'Ausilio et al., 2012; Schwartz et al., 2012).

Pris ensemble, ces modèles appuient ainsi l'existence de contraintes sensorielles et motrices dans le contrôle et la reconnaissance/catégorisation des actes de parole produite et perçue. Dans ce cadre, cette étude en imagerie par résonance magnétique fonctionnelle a pour but une exploration fine des réseaux neuronaux et des cartes sensorielles et motrices activés aussi bien lors de la production que de la perception des voyelles du Français. En regard des modèles présentés ci-dessus, l'utilisation des voyelles devait ainsi permettre de déterminer un possible couplage fonctionnel des systèmes de perception et de production de la parole et ce, pour des unités élémentaires de la parole, spécifiées et décrites en termes articulatoires et acoustiques précis autour d'un jeu universel de traits phonétiques et peu contaminées par des effets complexes de coarticulation et n'impliquant pas de traitements linguistiques morphologiques, sémantiques ou syntaxiques.

Codage des unités de parole

Nous l'avons vu, les modèles de perception de la parole divergent sur la localisation exacte des régions impliquées dans le décodage acoustico-phonétique et supposées représenter le centre des représentations phonétiques (région antérieure du gyrus/sulcus supérieur temporal vs région postérieure du gyrus/sulcus supérieur temporal et gyrus temporal moyen; par exemple, Scott et Johnsrude, 2003; Obleser et al., 2006; Rauschecker et Scott, 2009; Rauschecker, 2011 vs Belin et Zatorre 2000; Hickok et Poeppel, 2007). Certains modèles (neuro)computationnels de production de la parole impliquent quant à eux des cartes sensorimotrices détaillées dans lesquelles chaque articulateur a un rôle et une représentation spécifique et postulent un centre des représentations de parole (speech sound map) instancié dans le cortex prémoteur ventral et la partie postérieure du gyrus frontal inférieur (Guenther, Ghosh et Trouville, 2006; Guenther et Vladusich, 2012).

A la lumière des modèles neurobiologiques de perception et de production présentés ci-dessus, une question fondamentale est celle d'un possible codage distribué au sein des régions sensorielles et motrices des représentations de parole. Nous présenterons dans ce sous-chapitre les études portant sur le codage cérébral des unités de parole.

Tonotopie. Similairement à l'**arrangement somatotopique** des représentations motrices et somatosensorielles des différents articulateurs de la parole (voir Chapitre 1), une tonotopie spatiale des fréquences des sons perçus est observée au sein des récepteurs auditifs au niveau de la cochlée jusqu'à certaines régions auditives temporales. Ainsi Lauter et collègues (1985) ont observé pour la première fois une cartographie postéro-antérieure et médio-latérale pour les fréquences hautes-basses au sein du cortex auditif primaire humain en utilisant la méthode de tomographie par émission de positrons (TEP). Ces résultats démontrant une tonotopie au sein du cortex auditif primaire chez l'homme ont ensuite été répliqués en IRMf (Wessinger et al., 1997). Des études utilisant des enregistrements unicellulaires au moyen de microélectrodes ont également révélé des organisations postéro-antérieure dans le cortex auditif primaire et antéro-postérieure dans le cortex auditif secondaire pour les sons de fréquences hautes-basses chez le singe macaque (Merzenich et Brugge, 1973; Morel et al., 1993; Rauschecker et al., 1995).

Phonémotopie. A l'instar de la tonotopie, certaines études ont essayé de déterminer une possible topographie de codage des unités de parole et notamment des phonèmes (soit une "phonémotopie", expression empruntée à Obleser et collègues, 2003). Cette question a été d'abord abordée par des études MEG qui ont notamment observé au sein du cortex auditif postérieur des gradients de position selon les axes antéro-postérieur, latéro-médial et inféro-supérieur encodant les différents traits phonétiques des voyelles /i/, /u/ et /a/ (Obleser et al., 2003b; Shestakova et al., 2004; Schäringer et al., 2011 pour la langue turque). Une étude IRMf utilisant des analyses classiques univariées suggère également une organisation topographique du trait d'antériorité dans l'axe antéro-postérieur d'une région antérieure du gyrus temporal supérieur (Obleser et al., 2006). D'autres études ont aussi trouvé une organisation topographique des activations cérébrales évoquées lors de la production de syllabes au sein du cortex moteur primaire correspondant à l'organisation somatotopique des articulateurs orofaciaux impliqués dans la production de ces syllabes (par exemple /pa/-lèvres vs /ta/-langue; Pulvermüller et al., 2006; Skipper et al., 2007).

Codage distribué³⁷(analyses multi-variées de patterns multi-voxels, MVPA). Des études IRMf portant sur la perception de la parole et utilisant des analyses multivariées de reconnaissance statistique de patterns d'activation ont également démontré des classifications statistiquement significatives entre voyelles perçues (/i/, /a/, /u/) au niveau des aires temporales supérieures, depuis le gyrus latéral de Heschl (parties antérieure et postérieure) jusqu'au sulcus temporal supérieur (Formisano et al., 2008;

³⁷ MVPA, multi-voxel pattern analysis. Voir Norman et al., 2006 pour une revue sur l'application des méthodes d'analyse multivariée de patterns d'activation dans divers domaines des neurosciences cognitives (directions, catégorisation d'objets).

voir Figure 1.11) et entre voyelles et consonnes/syllabes au sein de la partie antérieure du gyrus temporal supérieur (Obleser et al., 2010). Néanmoins, ces études se sont spécifiquement focalisées sur le système auditif, et, à notre connaissance, aucune étude n'a étudié à ce jour le codage neural des traits phonétiques de voyelles aussi bien perçues que produites.

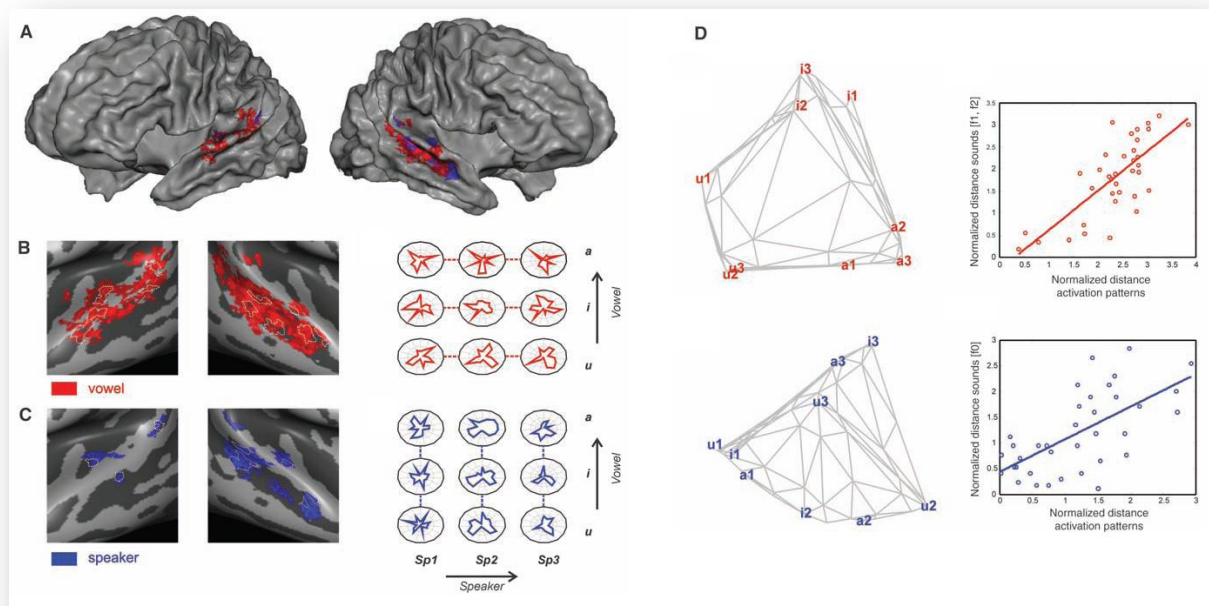


Figure 1.11 : Cartes corticales et empreintes topographiques neurales observées lors de la perception auditive de voyelles en fonction du type de voyelle perçu (en rouge) ou du locuteur (en bleu). A) Activations surfaciques moyennes. B) et C) Activités moyennes (gauche) et individuelles (droite) utilisées comme régions d'intérêt pour les analyses MVPA. D) Représentation topographique des voyelles (gauche) et corrélations entre distances spatiales et acoustiques de voyelles (droite) suite aux analyses MVPA en fonction du type de voyelle (haut) ou du locuteur (bas). (figures empruntées à Formisano et al., 2008)

Cadre de l'étude

La présente étude IRMf avait pour but de déterminer les corrélats neuronaux communs et distincts aux systèmes de perception et de production de voyelles du Français.

Face aux modèles neurobiologiques présentés dans la partie introductive, une question fondamentale est celle de l'implication commune d'une voie de traitement postéro-dorsale (Hickok et Poeppel, 2000, 2004, 2007; Scott et Johnsrude, 2003; Callan et al., 2004, 2010; Wilson et Iacoboni, 2006; Skipper et al., 2007; Rauschecker et Scott, 2009; Hickok, Houde et Rong, 2011; Rauschecker, 2011; Tremblay et Small, 2011), permettant l'association entre représentations auditives et motrices de la parole, lors de la perception et de la production des voyelles. Une seconde question concerne une possible modulation des réponses corticales sensorielles lors de la production des voyelles, telle qu'observée lors de la production de séquences de parole plus complexes (Paus et al., 1996; Numminen et Curio, 1999; Curio et al., 2000; Houde et al., 2002; Christoffels, Formisano et Schiller, 2007; Christoffels et al., 2011). Les interactions entre systèmes sensoriels et moteurs de la parole étant présumées dépendre de la complexité acoustique/articulatoire et phonétique des unités de parole perçues et/ou produites, de tels résultats observés pour des voyelles, considérées comme des unités élémentaires de la parole et peu contaminées par des effets complexes de coarticulation, apporteraient des arguments empiriques importants en faveur d'un couplage fonctionnel des systèmes de perception et de production de la parole chez l'adulte.

Un second objectif de cette étude était de déterminer un possible codage sensorimoteur distribué, soit une ségrégation topographique au niveau des cortex auditif et moteur, des voyelles et des traits

phonétiques d'antériorité, d'arrondissement et d'ouverture. Dans ce but, les participants ont d'abord écouté passivement 9 voyelles différentes maintenues du français (multiples occurrences des voyelles /i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/), propres à chaque sujet et enregistrées préalablement à l'étude. Dans la tâche suivante, réalisée lors de la même session IRM et en utilisant exactement les mêmes paramètres d'acquisition, les participants ont produit ouvertement ces mêmes voyelles. Ces 9 voyelles ont été sélectionnées dans le but de comparer les traits phonétiques d'ouverture (voyelles fermées, mi-fermées et mi-ouvertes), d'antériorité (voyelles antérieures et postérieures) et d'arrondissement (voyelles arrondies et non arrondies). De manière attendue (Schwartz et al., 1997a,b; Ladefoged, 2006; Ménard et al., 2008), les analyses acoustiques des enregistrements préalables à l'étude ont montré une corrélation positive entre le trait d'ouverture (correspondant à l'ouverture de la mâchoire) et la valeur du premier formant F_1 des voyelles fermées, mi-fermées et mi-ouvertes. De même, une corrélation négative a été observée entre les traits d'arrondissement et de postériorité (correspondant respectivement à l'arrondissement des lèvres et à la position rostro-caudale de la langue) avec la valeur du second formant (F_2), décroissant en passant des voyelles non arrondies antérieures aux voyelles arrondies antérieures puis aux voyelles arrondies postérieures. Par conséquent, notre design expérimental permettait de varier de manière orthogonale les deux dimensions d'ouverture (fermées, mi-fermées, mi-ouvertes) et de postériorité/arrondissement (antérieures non arrondies, antérieures arrondies, postérieures arrondies) pour les 9 voyelles sélectionnées³⁸.

Il est à noter que la plupart des études portant sur un possible couplage fonctionnel des systèmes de perception et de production de la parole ont utilisé lors des tâches perceptives et/ou motrices un protocole d'acquisition IRMf classique impliquant un bruit très important du scanner (Wilson et al., 2004; Okada et Hickok, 2006; Skipper et al., 2007; Callan et al., 2010; Tremblay et Small, 2011).

Comme mentionné précédemment, une altération de la qualité du feedback acoustique lors de la production de la parole induit non seulement des mécanismes de compensation articulatoire mais aussi des modifications en terme d'activation cérébrale en relation avec les processus de monitoring auditif et de contrôle moteur en ligne de la parole (Christoffels, Formisano et Schiller, 2007; Tourville, Reilly et Guenther, 2008; Zheng, Munhall et Johnsrude, 2010; Christoffels et al., 2011). De plus, certains modèles neurobiologiques de la perception de la parole proposent aussi que le système moteur serait plus fortement recruté en condition d'écoute difficile dans le but de permettre une meilleure reconnaissance/interprétation phonétique des sons de parole perçus (Callan et al., 2004; Skipper et al., 2005; Wilson et Iacoboni, 2006; Skipper et al., 2007). Cette hypothèse est appuyée par des études IRMf montrant une activité accrue de régions du système moteur inversement fonction du degré d'intelligibilité de la parole perçue, comme lors de parole masquée ou bruitée (Binder et al., 2004; Zekveld et al., 2006) ou encore lors de processus d'identification auditive de phonèmes non-natifs vs natifs (Callan et al., 2004; Wilson et Iacoboni, 2006). Pour cette raison, nous avons utilisé la technique d'acquisition de sparse sampling dans le but de réduire le bruit du scanner lors des tâches de perception et de production des voyelles ainsi que d'éliminer de possibles artefacts liés aux mouvements articulatoires lors de la tâche de production des voyelles.

³⁸ Du fait de l'utilisation de ce plan expérimental 3 x 3 (voyelles fermées, mi-fermées, mi-ouvertes x voyelles antérieures non arrondies, antérieures arrondies, postérieures arrondies), la voyelle ouverte antérieure non arrondie /a/ n'a pas été incluse dans cette étude.

1.2.2 Méthode

Protocole

Participants

Quatorze volontaires adultes sains ont participé à l'étude (11 hommes; âge : 21-44 ans, moyenne : 29 ans). Le recrutement des sujets est explicité dans la partie "Considérations Méthodologiques".

Stimuli

Les 9 voyelles (/i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/) ont été sélectionnées dans le but de comparer les traits phonétiques d'ouverture (voyelles fermées, mi-fermées et mi-ouvertes), d'antériorité (voyelles antérieures et postérieures) et d'arrondissement (voyelles arrondies et non arrondies). Préalablement à la session IRM, un premier entretien avait pour but d'expliquer à chaque participant le déroulement général de l'étude, les différentes tâches perceptives et motrices ainsi que de procéder à un enregistrement des voyelles dans une chambre sourde en vue de les utiliser pour la tâche de perception de parole.

Chaque voyelle a été produite 10 fois selon un ordre pseudo-aléatoire (de manière à ce qu'une même voyelle ou un même trait phonétique d'arrondissement, d'antériorité ou d'ouverture ne soit pas jamais répété). Pour chaque essai, une consigne visuelle ("ii", "ou", "éé" ...) indiquait au sujet la voyelle à produire.

Précédé d'une croix de fixation durant 500ms, l'affichage de cette consigne durant 1s servait également de repère temporel aux sujets pour la réalisation et la production maintenue de la voyelle. Il était demandé aux participants d'initier chacune de ces tâches depuis une position immobile de repos, les muscles linguaux et labiaux relâchés et de revenir à cette position initiale de repos à la fin du mouvement. L'intervalle entre chaque essai était de 5s. Les paramètres d'enregistrement des voyelles étant similaires à ceux utilisées lors de la session IRM, cette procédure servait donc également d'entraînement à la tâche de production dans le scanner.

En tout, 1260 voyelles (14 sujets x 9 voyelles x 10 productions) ont été enregistrées indépendamment avec un taux d'échantillonnage de 44.1 kHz et une résolution de 16 bits. Pour chaque voyelle et chaque participant, la fréquence fondamentale (F_0), le premier et second formant (F_1 et F_2) et la durée de production ont été calculés à l'aide du logiciel Praat (Institute of Phonetic Sciences, University of Amsterdam, NL; voir Figure 1.12 et Table 1.4). Six des occurrences les plus stables et naturelles ont été sélectionnées par voyelle et par volontaire en vue de servir à la tâche de perception.

La durée moyenne des voyelles était de 521ms (± 42 ms), avec un décalage de production de la voyelle par rapport à l'onset de la consigne visuelle d'environ 500ms. La distribution des valeurs formantiques des voyelles produites apparaît tout à fait concordante en regard de la littérature, malgré une variabilité importante due à des différences interindividuelles (notamment les différences entre locuteurs masculins et féminins).

De manière attendue (Schwartz et al., 1997a,b; Ladefoged, 2006; Ménard et al., 2008), les valeurs moyennes de F_1 des voyelles fermées, mi-fermées et mi-ouvertes étaient fonction du degré d'ouverture de la voyelle (en moyenne, 247Hz, 371Hz et 524Hz, respectivement). Parallèlement, les valeurs moyennes de F_2 des voyelles non arrondies antérieures, arrondies antérieures et arrondies postérieures étaient positivement corrélées avec le degré de (non)arrondissement et d'antériorité (en moyenne, 2024Hz, 1536Hz et 839Hz, respectivement).

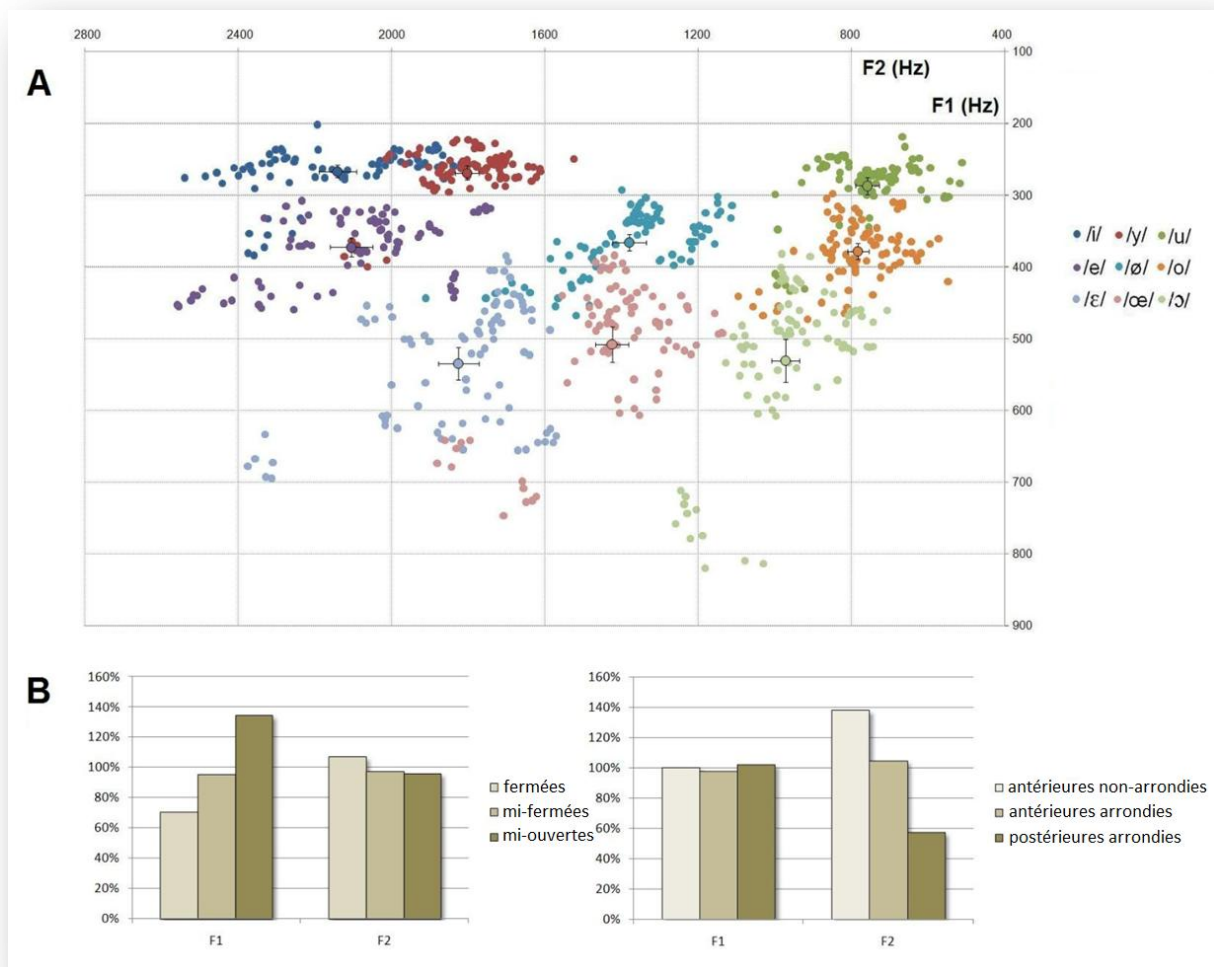


Figure 1.12 : A) Espace acoustique F_1/F_2 des 9 voyelles du Français enregistrées préalablement à l'étude IRMF pour chacun des sujets. Chaque point représente la valeur moyenne de six occurrences d'une même voyelle pour un seul sujet. Les valeurs moyennes pour chaque type de voyelle pour les 14 sujets sont également indiquées (les barres d'erreurs représentent l'erreur standard; voir Table 1.4). B) Valeurs formantiques moyennes F_1/F_2 des voyelles exprimées en pourcentage de la moyenne générale pour les voyelles fermées, mi-fermées et mi-ouvertes (gauche) et pour les voyelles antérieures non arrondies, antérieures arrondies et postérieures arrondies (droite).

Voyelle	Ouverture	Antériorité	Arrondissement	F_0	F_1	F_2	Durée
/i/	Fermée	Antérieure	Non arrondie	157 ±13	267 ±09	2141 ±49	0.521 ±0.041
/y/	Fermée	Antérieure	Arrondie	154 ±12	269 ±10	1803 ±31	0.506 ±0.044
/u/	Fermée	Postérieure	Arrondie	158 ±12	287 ±12	759 ±30	0.519 ±0.044
/e/	Mi-fermée	Antérieure	Non arrondie	152 ±12	372 ±13	2105 ±56	0.525 ±0.042
/ø/	Mi-fermée	Antérieure	Arrondie	149 ±11	365 ±11	1380 ±43	0.520 ±0.040
/o/	Mi-fermée	Postérieure	Arrondie	151 ±12	378 ±11	784 ±28	0.525 ±0.042
/ɛ/	Mi-ouverte	Antérieure	Non arrondie	148 ±12	534 ±23	1826 ±53	0.505 ±0.045
/œ/	Mi-ouverte	Antérieure	Arrondie	149 ±12	508 ±24	1425 ±44	0.524 ±0.042
/ɔ/	Mi-ouverte	Postérieure	Arrondie	149 ±12	530 ±29	973 ±36	0.541 ±0.041

Table 1.4 : Valeurs moyennes de la fréquence fondamentale (F_0 en Hz), des premier et second formants (F_1 , F_2 en Hz) et de la durée (en ms) pour toutes les voyelles préenregistrées auprès des participants et utilisées comme stimuli dans la tâche de perception. Les erreurs standards sont indiquées.

Procédure expérimentale

La passation IRM consistait en 2 tâches d'écoute passive et de production ouverte des voyelles /i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/. En vue de minimiser d'éventuels effets de simulation motrice, la tâche de perception auditive a toujours précédé la tâche de production.

Perception. La tâche de perception a été réalisée lors des trois premiers runs fonctionnels (voir partie “Considérations Méthodologiques”, Figure 1.1). Elle consistait en l'écoute passive d'une des 9 voyelles maintenues du Français (/i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/) propres à chaque sujet et enregistrées préalablement à l'étude (6 occurrences différentes par voyelle ont été présentées). Une condition de repos, sans aucun mouvement ni stimulation auditive, servait de condition contrôle. Durant l'expérience, l'écoute d'une voyelle ou une condition de repos apparaissait toutes les 10 secondes (du fait de l'acquisition des images fonctionnelles de type sparse sampling) selon un ordre pseudo-aléatoire (une même voyelle ou un même trait phonétique ne pouvant survenir plus de deux fois de suite). Chaque condition consistait en 18 essais. Concernant le protocole d'acquisition en sparse sampling, l'intervalle de temps séparant la perception d'une voyelle et le milieu d'acquisition de l'image fonctionnelle correspondante était contrebalancé entre chaque essai et était de 4s, 5s ou 6s. 180 scans fonctionnels ont ainsi été acquis ((9 voyelles + 1 condition de repos) x 18 scans). Lors de chaque run fonctionnel, 3 scans ont été ajoutés en début de session pour équilibrer le signal IRM et ont ensuite été supprimés des analyses.

Production. La tâche de production était réalisée lors des 3 runs fonctionnels suivants (voir partie “Considérations Méthodologiques”, Figure 1.1). Il était demandé aux participants de produire ouvertement chacune des 9 voyelles de manière maintenue. Tout comme lors de la tâche de perception, une condition contrôle, sans mouvement ni stimulation auditive, a été ajoutée. Le sujet devait produire l'une des 9 voyelles ou la condition de repos toutes les 10 secondes (du fait de l'acquisition des images fonctionnelles de type sparse sampling) selon un ordre pseudo-aléatoire (une même voyelle ou un même trait phonétique ne pouvant survenir plus de deux fois de suite). Chaque condition consistait en 18 essais. Comme pour la tâche de perception, l'intervalle de temps séparant la production d'une voyelle et le milieu d'acquisition de l'image fonctionnelle correspondante était contrebalancé entre chaque essai et était de 4s, 5s ou 6s. Pour chaque essai, une consigne visuelle (“ii”, “ou”, “repos”...) indiquait au sujet la condition à réaliser. Précédée d'une croix de fixation durant 500ms, l'affichage de cette consigne durant 1s servait également de repère temporel aux sujets pour la réalisation de la tâche. Il était demandé aux participants d'initier chacune de ces tâches depuis une position immobile de repos, les muscles linguaux et labiaux relâchés et de revenir à cette position initiale de repos à la fin du mouvement. Un enregistrement des productions du sujet a été réalisé de manière à contrôler *a posteriori* d'éventuelles erreurs de production et de les supprimer des analyses suivantes. 180 scans fonctionnels ont ainsi été acquis ((9 voyelles + 1 condition de repos) x 18 scans). Lors de chaque run fonctionnel, 3 scans ont été ajoutés en début de session pour équilibrer le signal IRM et ont ensuite été supprimés des analyses.

Prétraitements

Préalablement aux analyses statistiques, une série de prétraitements des images fonctionnelles a été réalisée avec le logiciel SPM 5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) sous environnement Matlab 7.1 (Mathworks, Natick, MA). Ces prétraitements incluaient des procédures de réalignement, de normalisation dans l'espace commun du Montreal Neurological Institute (repère MNI) et de lissage spatial (voir partie “Considérations Méthodologiques” pour une description détaillée de ces procédures).

Erreurs de production. L'écoute et l'inspection du signal acoustique de tous les essais (incluant 2268 voyelles produites et 252 essais de repos) ont été effectuées pour s'assurer de l'exactitude des réponses des participants. Trois types d'erreurs ont été observés : des omissions, des productions

erronées ou des hésitations. En tout, 19 images fonctionnelles relatives à ces erreurs de productions ont été écartées des analyses (en moyenne, $0.75\% \pm 0.25\%$). Cette analyse démontre la réalisation correcte de plus de 99% des voyelles produites (aucun participant n'ayant commis plus de 10% d'erreurs).

Analyses statistiques

Suite aux prétraitements, les données fonctionnelles ont été analysées avec le logiciel SPM 5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) sous environnement Matlab 7.1 (Mathworks, Natick, MA). Les régions cérébrales activées ont été labellisées en utilisant la toolbox SPM Anatomy (Eickhoff et al., 2005) et, si nécessaire, en utilisant le logiciel Talairach Daemon (Lancaster et al., 2000). Les analyses par région d'intérêt (ROI) ont été effectuées via la toolbox SPM Anatomy. Pour la visualisation des résultats, les cartes d'activation ont été superposées sur un template anatomique standard en utilisant le logiciel MRICRON (<http://www.sph.sc.edu/comd/rorden/mricron/>).

Analyses individuelles

Au niveau individuel et suite aux prétraitements des images fonctionnelles (voir partie "Considérations Méthodologiques"), les activations neuronales liées à la perception et production des voyelles ont été estimées par l'utilisation d'un Modèle Linéaire Général (GLM; Friston et al., 1995). Pour chaque participant, un second GLM a été estimé de manière à évaluer les activités reliées aux 3 délais d'acquisition (4s, 5s et 6s) lors des tâches de perception et de production.

Les modèles incluaient des régresseurs d'intérêt reliés soit aux 2 tâches perceptives et motrices pour chacune des 9 voyelles (2 x 9 régresseurs représentés chacun par 18 images fonctionnelles) indépendamment des délais d'acquisition, soit aux 3 délais d'acquisition pour les deux tâches motrices et perceptives (2 x 3 régresseurs représentés chacun par 54 images fonctionnelles) et des régresseurs de non-intérêt liés aux paramètres de réalignement. Les conditions de repos formaient une ligne de base (2 x 18 images fonctionnelles).

Pour les deux modèles, la réponse de type hémodynamique associée à chaque événement a été modélisée par une fonction à réponse impulsionnelle finie et à impulsion unique (FIR) pour chaque scan fonctionnel. Avant l'estimation du modèle, un filtrage des basses fréquences *a priori* non reliées aux conditions expérimentales (variations lentes d'origine physiologique) a été appliqué (passe-haut de fréquence de coupure de 1/128 Hz). Les estimations des réponses FIR modélisées ont été calculées de manière à correspondre avec le décours temporel du signal BOLD observé pour chaque voxel et chaque condition. Pour chaque participant, des cartes d'analyse statistique individuelles ont ainsi été calculées pour chaque voyelle perçue ou produite dans le premier modèle et pour chacun des 3 délais d'acquisition pour les deux tâches motrices et perceptives dans le second modèle, par contraste avec la condition contrôle. Ces cartes statistiques individuelles ont été ensuite utilisées pour les analyses de groupe.

Analyses de groupe

Afin de permettre une généralisation des résultats observés pour la population étudiée (Friston, Holmes et Worsley, 1999), des analyses de groupe de second niveau à effet aléatoire ont été réalisées au moyen d'analyses de variance (ANOVA) à mesures répétées.

Perception et production des voyelles

De manière à déterminer les activités cérébrales liées à la perception et la production de voyelles, une ANOVA à mesures répétées a été réalisée avec comme facteurs intra-sujet la tâche (2 niveaux, perception ou production) et la voyelle perçue ou produite (9 niveaux relatifs à chacune des voyelles). La variable “sujets” a été prise en compte comme facteur aléatoire.

Réseaux généraux : Deux contrastes t unidirectionnels ont été calculés afin de déterminer les régions cérébrales activées de manière spécifique lors de la perception et lors de la production des voyelles (toutes voyelles confondues) par rapport à la condition de repos. Dans le but d’identifier les activations communes/partagées aux 2 tâches, une analyse de conjonction (Friston et al., 1999; Nichols et al., 2005) a également été effectuée. Finalement, un contraste F bidirectionnel correspondant à l’effet principal de la tâche a été calculé dans le but de déterminer les différences d’activations cérébrales entre les deux tâches.

Traits phonétiques : Séparément pour les tâches de perception et de production, des analyses supplémentaires ont été réalisées en regard du trait d’ouverture, d’une part, et des traits d’arrondissement et d’antériorité, d’autre part. Ainsi, aussi bien pour la tâche de perception que pour celle de production, des contrastes t ont été calculés pour les voyelles fermées, mi-fermées et mi-ouvertes et pour les voyelles antérieures non arrondies, antérieures arrondies et postérieures arrondies. Enfin, des contrastes F bidirectionnels correspondant à l’effet principal du trait d’ouverture et du trait d’arrondissement/antériorité ont été calculés dans le but de déterminer les différences d’activations cérébrales entre types de voyelles.

Voyelles : Des contrastes t ont été enfin calculés pour chacune des 9 voyelles, relativement à la condition de repos, lors des deux tâches de perception et de production. Des contrastes F bidirectionnels correspondant à l’effet principal de la voyelle ont été calculés dans le but de déterminer les différences d’activations cérébrales entre les 9 voyelles et ce, lors de la tâche de perception et lors de la tâche de production.

Délais d’acquisition

Une seconde analyse de groupe a été réalisée dans le but de comparer les activations cérébrales fonctions du délai d’acquisition des images fonctionnelles lors des 2 tâches de perception et de production, indépendamment du type de voyelles perçues ou produites. En plus du décours temporel des activations, cette analyse devait ainsi permettre de vérifier l’adéquation des délais d’acquisition choisis. Pour ce faire, une seconde ANOVA à mesures répétées a été effectuée avec comme facteurs intra-sujet la tâche (perception ou production) et le délai d’acquisition (incluant 3 niveaux liés aux délais d’acquisition de 4s, 5s et 6s par rapport à la réalisation de la tâche motrice). La variable “sujets” a été prise en compte comme facteur aléatoire.

Pour chaque tâche, trois contrastes t ont été calculés de manière à déterminer les régions cérébrales activées de manière spécifique pour chacun des trois délais d’acquisition par rapport à la ligne de base. Dans le but d’identifier des activations communes et distinctes aux 3 délais d’acquisition, des analyses de conjonction et d’effet principal du délai d’acquisition ont été calculées pour chaque tâche. Enfin, de manière à déterminer des décours temporels d’activation différents entre tâches relativement aux 3 délais d’acquisition, l’interaction entre délais d’acquisition et tâche a également été estimée.

1.2.3 Résultats

Les résultats des analyses de groupe sont reportés selon un seuil statistique défini à $p < .05$ après correction de type “False Discovery Rate” (FDR; Genovese et al., 2002) et une taille minimale des

clusters de 25 voxels. Pour toutes les analyses et contrastes, les pics d'activation maximum ont été déterminés dans chaque cluster, leur localisation a ensuite été labellisée avec la boîte à outils Anatomy de SPM (Eickhoff et al., 2005). Si une région n'avait pas pu être assignée avec Anatomy, elle a été déterminée avec le logiciel Talairach Daemon (Lancaster et al., 2000) grâce aux coordonnées du pic d'activation converties de l'espace MNI à l'espace stéréotaxique standard de Talairach et Tournoux (1988).

Perception et production des voyelles

La Table 1.5 répertorie les activations principales pour les tâches de perception et de production (par rapport à la ligne de base) ainsi que les activations communes révélées par l'analyse de conjonction. Les cartes d'activation sont illustrées dans la Figure 1.13.

Perception

Le réseau global de perception des voyelles implique de fortes activations auditives bilatérales au sein du gyrus/sulcus temporal supérieur (incluant les cortex auditifs primaires, secondaires et associatifs, depuis l'aire temporopolaire à l'extrémité antérieure du gyrus temporal supérieur jusqu'à la jonction temporo-pariétale à son extrémité postérieure) et s'étendant de manière dorso-médiale jusqu'au cortex insulaire postérieur et de manière ventrale jusqu'au gyrus temporal moyen. Des activations frontales ont été également observées, principalement dans l'hémisphère gauche, dans les régions triangulaire et operculaire du gyrus frontal inférieur et le cortex prémoteur ventral adjacent. De plus, des activations ont été trouvées au niveau du cortex préfrontal dorsolatéral gauche, du gyrus frontal supérieur gauche et du gyrus frontal moyen droit. Des activations des lobules pariétaux supérieur et inférieur sont observées dans le précuneus, le gyrus supramarginal et l'opercule rolandique. D'autres activations ont été trouvées au sein des ganglions de la base (noyau caudé et substance noire), dans les structures limbiques (thalamus, cortex cingulaire et hippocampe) et dans le noyau rouge. Enfin des activations spatialement circonscrites des aires visuelles striées et extrastriées ont également été observées.

Production

En plus de l'activation des cortex auditifs telle qu'observée pour la tâche de perception, le réseau global de production des voyelles implique des activations bilatérales des cortex prémoteur et sensorimoteur primaire, du gyrus frontal inférieur postérieur (pars opercularis) et de l'opercule rolandique. Des activations supplémentaires sont observées dans l'aire motrice supplémentaire (SMA-proper), s'étendant jusqu'au cortex cingulaire antérieur, dans le cortex insulaire gauche, les lobules pariétaux supérieur et inférieur (précunéus, opercule pariétal et gyrus supramarginal), les ganglions de la base (noyau caudé et substance noire), le thalamus, le noyau rouge, le cervelet (culmen) et le cortex visuel strié et extrastrié.

Conjonction

L'analyse de conjonction révèle des activations bilatérales auditives communes aux tâches de perception et de production des voyelles (incluant les cortex auditifs primaire, secondaire et associatif) et s'étendant caudalement à l'aire temporopolaire gauche, dorsalement au gyrus supramarginal gauche et à l'opercule pariétal droit. Les activations de la partie operculaire du gyrus frontal inférieur et du cortex prémoteur ventral adjacent sont observées dans l'hémisphère gauche. Enfin, des activations additionnelles ont été observées dans le thalamus, la substance noire, le noyau rouge et le cortex visuel strié et extra strié.

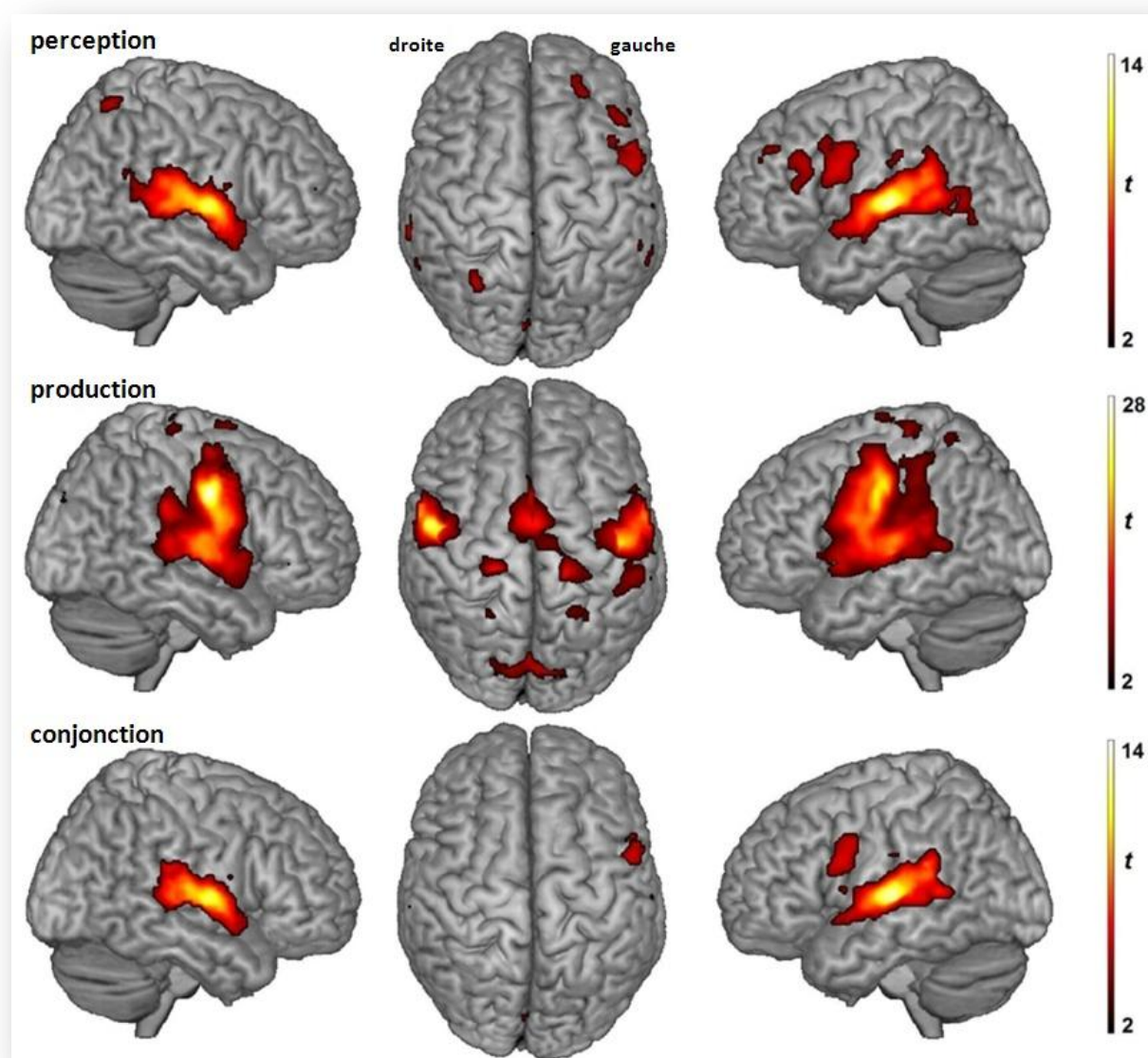


Figure 1.13 : Vue surfacique des régions cérébrales spécifiquement activées lors des tâches de perception et de production des voyelles et activations communes aux deux tâches révélées par l'analyse de conjonction. Résultats reportés à un seuil corrigé (FDR, $p < .05$) et une taille minimale des clusters de 25 voxels.

Régions	H	BA	Perception				BA	Production				BA	Conjonction			
			x	y	z	T		x	y	z	T		x	y	Z	T
Régions frontales																
Gyrus frontal inférieur	L	44	-54	8	26	5.87	44	-54	8	26	7.14	44	-54	8	26	4.87
	L	45	-52	32	16	3.17										
	R	44	60	4	16	3.07	44	54	4	-2	8.66					
Cortex prémoteur	L						6	-60	0	22	15.51					
	R	6	36	-16	28	3.59	6	46	-6	58	8.76					
Cortex moteur primaire	L						4	-46	-12	34	23.45					
	R						4	52	-6	36	27.48					
Gyrus frontal moyen	R	10	28	48	16	3.45										
Cortex préfrontal dorsolatéral	L	9	-42	8	30	5.32										
	L	46	-44	26	28	4.08										
Gyrus frontal supérieur	L	9	-28	42	34	3.49										
Aire motrice supplémentaire	L						6	-2	-4	62	11.27					
Régions temporales																
Aire temporopolaire	L	38	-56	8	-8	4.33	38	-56	8	-6	4.16	38	-56	8	-6	4.16
	R	38	38	2	-18	3.10										
Gvrus temporal supérieur	L	22	-60	-44	10	7.22	22	-58	-44	8	5.92	22	-58	-44	8	5.92

Gyrus temporal transverse	R	22	56	-8	4	12.64	22	50	-18	6	17.05	22	56	-8	4	12.64
	L	41	-52	-20	6	14.47	41	-36	-34	16	18.18	41	-52	-20	6	14.47
	R						41	42	-24	12	13.50	41	56	-26	10	7.64
Gyrus temporal moyen	L	21	-46	-60	10	3.59										
	L	37	-56	-58	14	3.33										
	R	21	40	-8	-10	3.63										
Gyrus fusiforme	R	37	42	-50	-14	2.94										
Régions pariétales																
Cortex somatosensoriel	L						3	-18	-40	50	2.57					
	R						2	26	-40	50	2.48					
Gyrus supramarginal	L	40	-48	-38	18	8.31	40	-52	-38	16	9.03	40	-48	-38	18	8.31
	R	40	54	-30	18	7.68	40	52	-30	28	10.17					
Opércule rolandique	L						43	-62	-8	8	14.27					
	R	43	62	-20	12	9.46	43	40	-8	18	9.35	43	62	-18	14	7.99
Lobule pariétal supérieur / Précuneus	L	7	-8	-74	36	3.49	7	-22	-50	42	4.47					
	R	7	30	-60	60	4.11	7	22	-54	64	3.04					
Insula																
Clastrum	L	13	-46	-20	26	4.20	13	-38	4	4	8.43					
	L		-32	2	8	3.48										
Système limbique																
Hippocampe	L		-28	-32	-2	3.89										
Gyrus parahippocampique	R	19	40	-48	-6	3.25										
Corps mammillaires	L		-2	-14	-12	3.68										
Cortex cingulaire antérieur	L	24	0	-8	26	4.78	32	-4	12	40	8.31					
	R	24	14	-6	32	3.96	32	8	14	34	4.22					
Cortex cingulaire postérieur	L	23	-6	-32	24	5.84										
Thalamus	L		-14	-32	-2	3.40		-12	-18	4	8.76		-14	-32	-2	3.40
	R		14	-14	-2	3.62		14	-14	-2	4.27		14	-14	-2	3.62
Ganglions de la base																
Striatum	L		-8	0	26	3.72		-28	-14	-2	9.89					
	R		8	16	16	4.76										
Substance noire	L		-12	-22	-12	3.48		-12	-22	-12	3.29		-12	-22	-12	3.29
	R		12	-26	-8	5.56		10	-24	-8	4.64		10	-24	-8	4.57
Tronc cérébral																
Noyau rouge	L		-8	-26	-6	4.23		-8	-26	-6	5.12		-8	-26	-6	4.23
Cervelet																
Culmen	L							-14	-64	-12	8.09					
Régions occipitales																
Cortex strié	L	17	-2	-70	12	3.62	17	-6	-74	12	10.54	17	-2	-70	12	3.62
	R	17	6	-84	10	4.59	17	20	-62	10	9.45					
Cortex extra-strié	L	19	-44	-64	-2	3.18						18	-2	-68	20	3.91
	R						18	10	-66	14	8.09					

Table 1.5 : Coordonnées dans l'espace MNI des pics d'activation maximale observés pour les tâches de perception et de production des voyelles et communs à ces 2 tâches (analyse de conjonction). Résultats reportés à un seuil corrigé (FDR, $p < .05$) et une taille minimale des clusters de 25 voxels.

Effet principal de la tâche

Activations – Production > Perception: Par rapport aux conditions de repos, l'effet principal de la tâche révèle des activations significatives lors de la tâche de production et quasi nulles lors de la tâche de perception dans les cortex moteur primaire et prémoteur, la partie caudale de l'aire motrice supplémentaire (SMA-proper), le gyrus supramarginal droit, le cortex insulaire, les ganglions de la base (noyau caudé et globus pallidus), certaines régions du système limbique (cortex cingulaire antérieur et thalamus) et du cortex visuel. Des activations de régions des gyri temporaux supérieur et transverse gauche et de l'opercule rolandique droit sont également observées lors des deux tâches mais avec une activité accrue lors de la tâche de production (voir Figure 1.14 et Table 1.6).

Déactivations : Des activations plus faibles que celles inhérentes à la condition de repos ('déactivations') ont été observée lors de la tâche de production dans la partie triangulaire du gyrus frontal inférieur gauche, dans le cortex préfrontal droit, le gyrus frontal moyen gauche, le cortex dorsolatéral préfrontal, le gyrus frontal supérieur et le gyrus frontal moyen gauche. Des

déactivations ont été observées dans la partie postérieure des gyri temporaux supérieur et moyen, la partie postérieure du gyrus supramarginal et le gyrus angulaire gauche adjacent, le lobule pariétal supérieur et le cortex cingulaire postérieur (voir Figure 1.14 et Table 1.6).

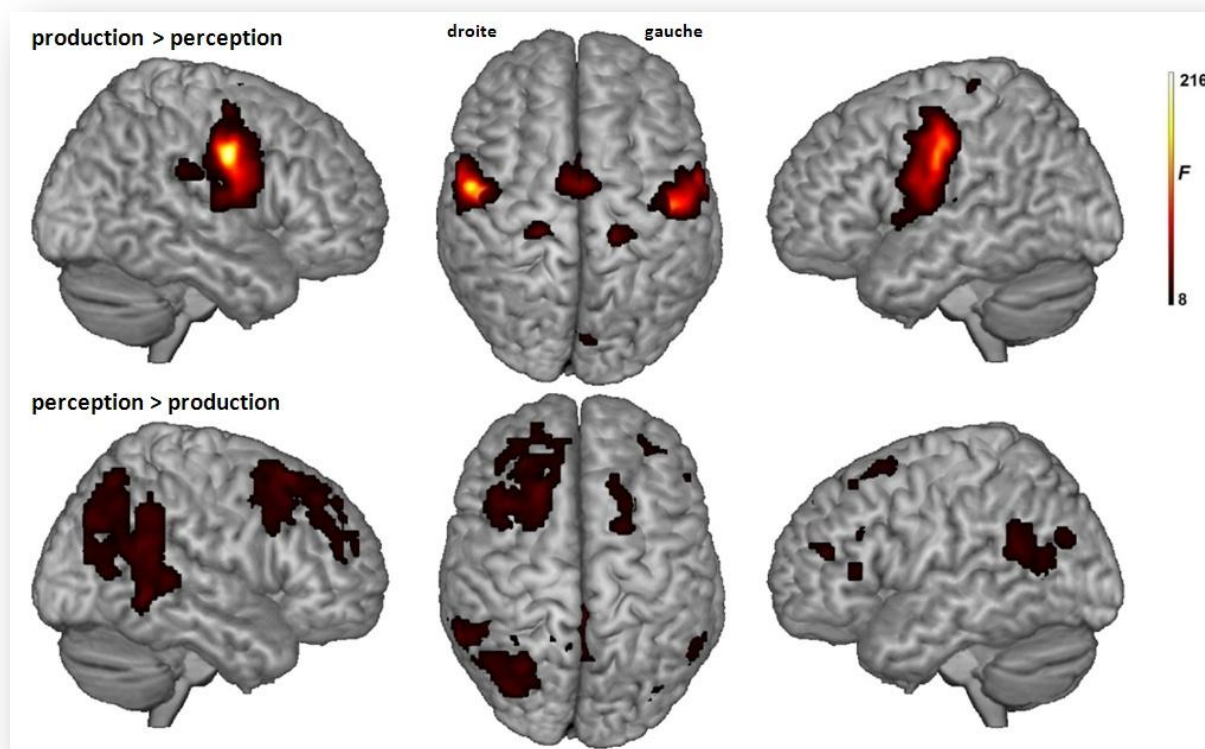


Figure 1.14 : Vue surfacique des régions cérébrales différenciellement activées lors des tâches de perception et de production des voyelles (FDR, $p < .05$ et taille minimale des clusters de 25 voxels).

Régions	H	BA	Effet principal - Tâche				Contrastes estimés	
			x	y	z	F	Perception	Production
Production > Perception								
Régions frontales								
Cortex prémoteur	L	6	-46	-12	34	171	-0.06	0.76
	R	6	52	-6	36	216	-0.03	0.71
Cortex moteur primaire	L	4	-40	-16	38	143	-0.05	0.73
	R	4	20	-30	60	36	-0.03	0.20
Aire motrice supplémentaire	L	6	-2	-4	62	39	-0.04	0.48
Régions temporales								
Gyrus temporal supérieur	L	22	-60	4	4	15	0.09	0.45
Gyrus temporal transverse	L	41	-42	-28	14	20	0.29	0.59
Régions pariétales								
Gyrus supramarginal	R	40	46	-28	26	19	0.01	0.21
Opercule rolandique	R	43	56	-10	14	31	0.13	0.47
Insular cortex								
Insula	L	13	-36	-10	16	16	-0.03	0.13
	R	13	38	-8	18	27	-0.03	0.19
Système limbique								
Cortex cingulaire antérieur	L	32	-4	10	40	19	-0.05	0.08
	R	32	20	32	16	16	-0.05	0.08
Thalamus	L		-12	-16	2	24	-0.02	0.15
	R		28	-2	-8	19	-0.04	0.16
Ganglions de la base								
Noyau caudé	R		36	-40	6	10	-0.04	0.07
Globus pallidus	L		-18	0	-2	13	-0.05	0.06

Régions occipitales								
Cortex strié	L	17	-14	-66	6	22	-0.01	0.30
	R	17	22	-62	8	27	-0.03	0.26
Cortex extra-strié	L	18	-12	-52	-4	16	-0.03	0.27
	L	19	-4	-82	30	12	0.02	0.29
	R	18	10	-62	0	16	-0.04	0.20
Production < Perception								
Régions frontales								
Gyrus frontal inférieur / cortex préfrontal	L	45	-52	30	10	13	0.07	-0.14
	R	9	34	10	36	18	0.03	-0.15
Gyrus frontal moyen	L	10	-38	46	18	23	0.84	-0.19
	R	6	42	24	46	20	0.00	-0.22
Cortex préfrontal dorsolatéral / Gyrus frontal supérieur	L	46	-40	18	28	10	0.06	-0.08
	L	8	-16	28	52	11	0.08	-0.09
	R	6	20	24	58	28	-0.02	-0.26
Gyrus frontal médial	R	8	38	18	54	23	0.01	-0.23
	L	6	-20	6	56	16	0.08	-0.12
Régions temporales								
Gyrus temporal supérieur	R	22	54	-54	16	20	0.01	-0.20
Gyrus temporal moyen	L	37	-38	-64	8	17	0.07	-0.12
	R	22	52	-48	-4	28	0.00	-0.19
	R	37	32	-38	-14	15	0.00	-0.14
Régions pariétales								
Gyrus angulaire	L	39	-34	-74	28	13	0.03	-0.16
Gyrus supramarginal	L	40	-54	-52	22	21	-0.01	-0.22
	R	40	60	-52	22	24	0.09	-0.17
Lobule pariétal supérieur / Précuneus	L	7	-6	-58	50	10	0.01	-0.16
	R	7	32	-66	54	21	0.02	-0.26
Système limbique								
Cortex cingulaire postérieur	L	31	-12	-42	32	12	0.03	-0.12
	R	31	4	-42	38	36	0.06	-0.28

Table 1.6 : Coordonnées dans l'espace MNI des régions différenciellement activées lors des tâches de perception et de production des voyelles (FDR, $p < .05$ et taille minimale des clusters de 25 voxels).

Effets principaux des degrés d'ouverture et d'arrondissement/antériorité

Ouverture : Lors de la tâche de production, l'effet principal du degré d'ouverture montre une activation accrue pour les voyelles mi-ouvertes, par rapport aux voyelles fermées et mi-fermées, dans la partie operculaire du gyrus frontal inférieur gauche, dans le cortex insulaire antérieur gauche, le gyrus supramarginal gauche, le cortex somatosensoriel droit et le cortex prémoteur et l'aire motrice supplémentaire, bilatéralement (voir Figure 1.15 et Table 1.7). Aucune région cérébrale n'a montré de différences significatives entre voyelles selon le degré d'ouverture lors de la tâche de perception.

Régions	H	BA	Effet principal				Contrastes estimés		
			x	y	z	F	fermées	mi-fermées	mi-ouvertes
Régions frontales									
Gyrus frontal inférieur	L	44	-52	8	6	11.08	0.51	0.55	1.02
Cortex prémoteur	L	6	-44	0	42	13.91	-0.03	0.09	0.38
	R	6	28	-8	54	11.83	-0.10	0.05	0.24
Aire motrice supplémentaire	L	6	-2	12	52	22.57	-0.17	0.18	0.70
Régions pariétales									
Lobule pariétal supérieur / Gyrus supramarginal	L	40	-32	-56	40	10.54	-0.17	-0.05	0.20
Cortex somatosensoriel	R	2	-44	-40	46	9.53	-0.09	-0.10	0.25
Cortex insulaire									
Insula	L	13	-30	20	10	12.50	0.04	0.22	0.44

Table 1.7 : Coordonnées dans l'espace MNI des régions différenciellement activées lors de la tâche de production en fonction du degré d'ouverture (FDR, $p < .05$ et taille minimale des clusters de 25 voxels).

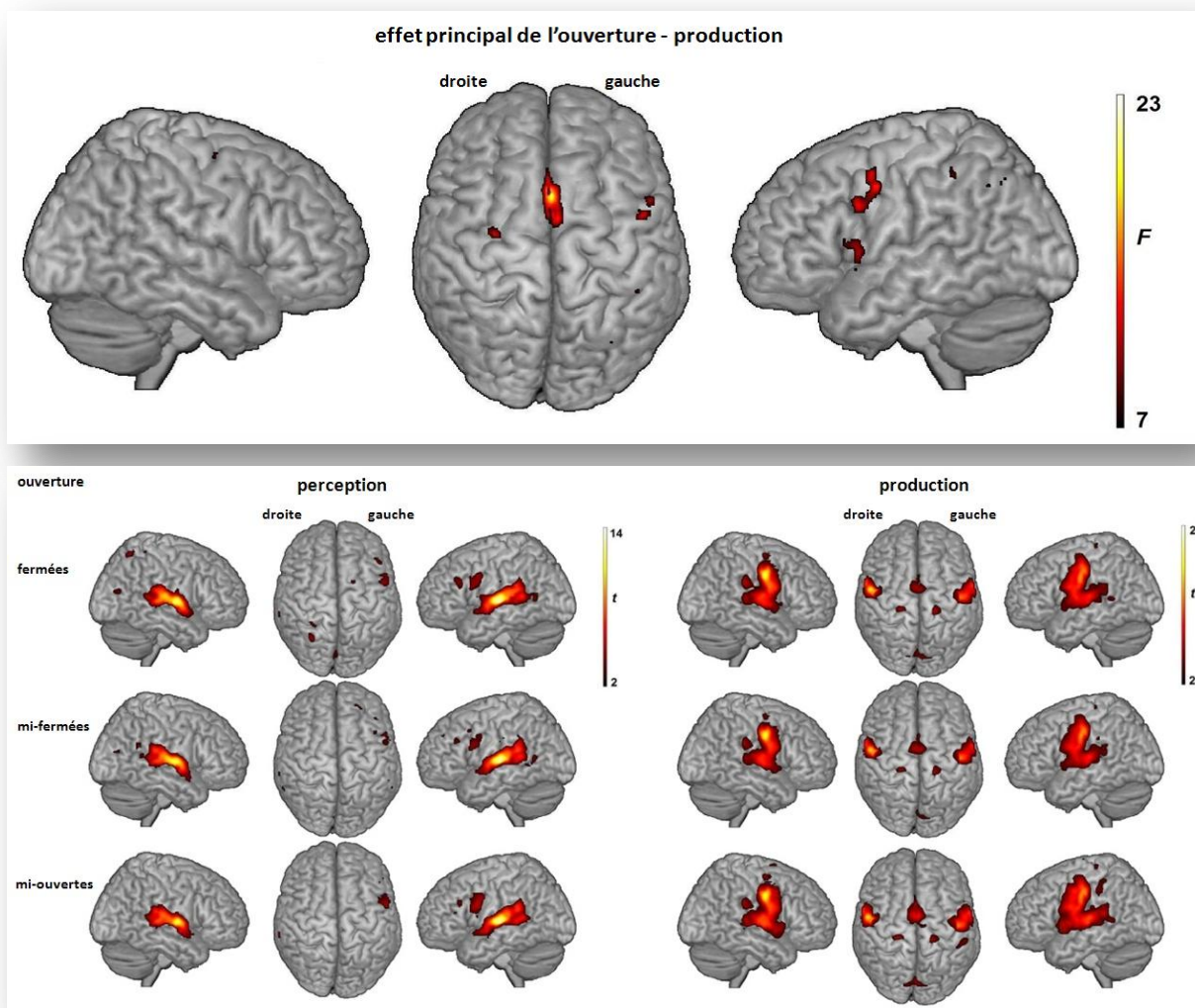


Figure 1.15 : (Haut) Régions cérébrales activées lors des tâches de perception et de production pour les voyelles fermées, mi-fermées et mi-ouvertes. (Bas) Régions différenciellement activées lors de la tâche de production en fonction du degré d'ouverture (FDR, $p < .05$ et taille minimale des clusters de 25 voxels).

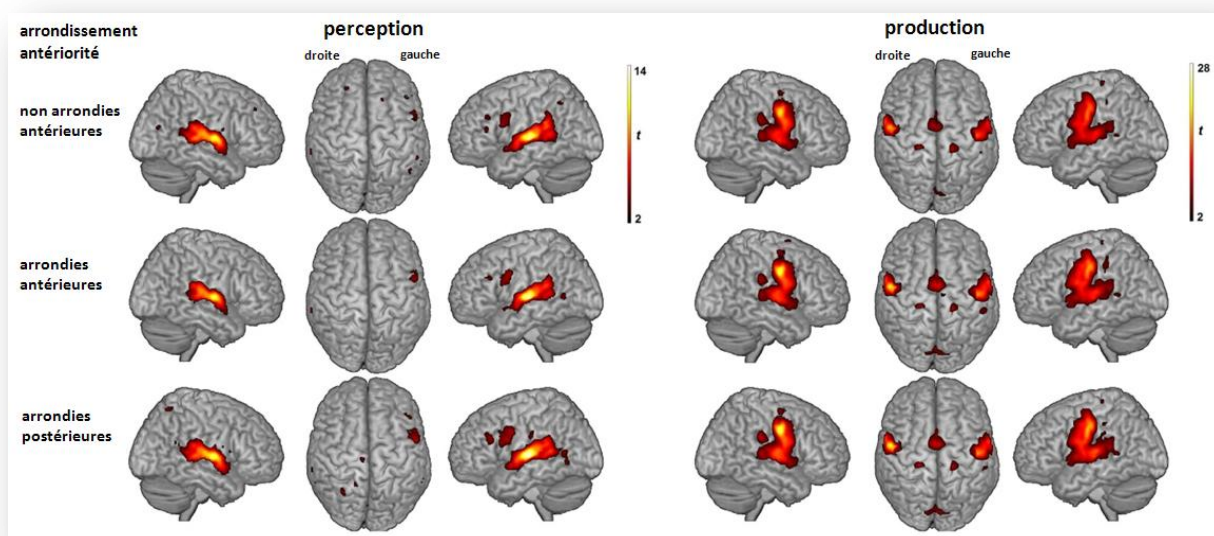


Figure 1.16 : Régions cérébrales activées lors des tâches de perception et de production pour les voyelles antérieures non arrondies, antérieures arrondies et postérieures arrondies (FDR, $p < .05$ et taille minimale des clusters de 25 voxels).

Arrondissement/antériorité : Aucune région cérébrale n'a montré de différences significatives entre voyelles selon le degré d'arrondissement/d'antériorité lors des tâches de perception et de production (voir Figure 1.16).

Effets principaux de la voyelle

Perception : Aucune différence significative d'activation entre voyelles n'a été trouvée lors de la tâche de perception.

Production : Pour la tâche de production, l'effet principal de la voyelle révèle des activations accrues pour les voyelles mi-ouvertes (particulièrement les voyelles arrondies) par rapport à toutes les autres voyelles dans l'aire motrice supplémentaire, les cortex frontal moyen, prémoteur et frontal inférieur, l'insula, le gyrus temporal supérieur et le gyrus supramarginal et ce, préférentiellement dans l'hémisphère gauche. Des activations accrues pour les voyelles mi-ouvertes sont également constatées dans l'hémisphère droit dans le précuneus, le thalamus, le cervelet (culmen), le gyrus cingulaire postérieur et les cortex visuels strié et extrastrié (voir Figure 1.17).

Activités cérébrales liées aux délais d'acquisition

Perception

Chaque délai d'acquisition des images fonctionnelles a été individuellement contrasté par rapport à la ligne de base. L'analyse de groupe révèle des activations significatives largement communes aux 3 délais d'acquisitions (4s, 5s, 6s), confirmant par là même la bonne adéquation de ces délais dans l'acquisition du signal BOLD lors de la tâche de perception. Néanmoins, l'effet principal des trois délais d'acquisition montre également des décours temporels significativement différents dans certaines régions corticales et sous-corticales (voir Table 1.8 et Figure 1.18).

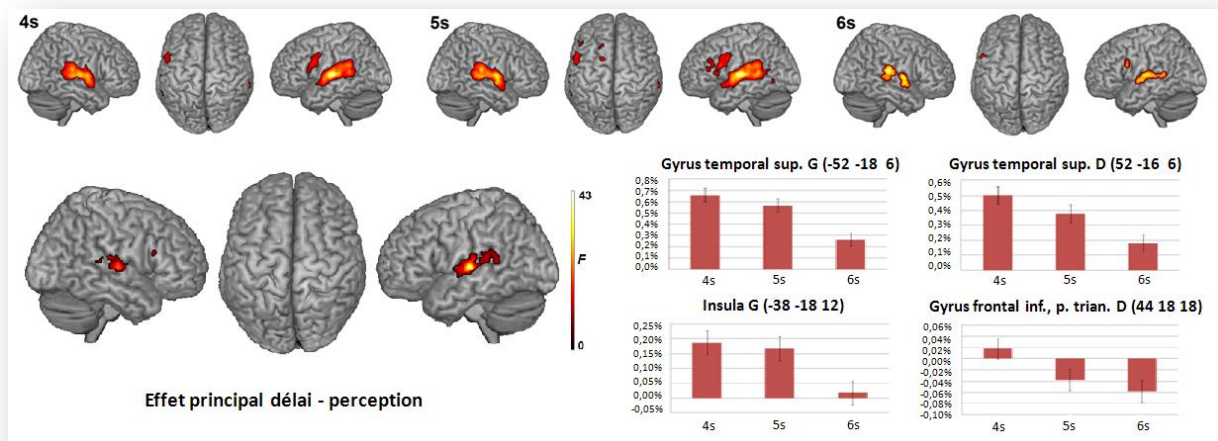


Figure 1.18 : Vue surfacique des régions cérébrales activées lors de la tâche de perception selon les 3 délais d'acquisition de 4s, 5s et 6s (haut) et effet principal (bas) montrant les régions cérébrales différemment activées selon les 3 délais d'acquisition. Les diagrammes représentent les valeurs des activations pour les trois délais d'acquisition dans certaines régions cérébrales indiquées dans l'effet principal ($p < .05$, FDR, et taille minimale des clusters de 25 voxels).

- Une courbe en cloche de l'activité neuronale au travers des trois délais d'acquisitions (pic d'activation maximal à 5s et significativement supérieur à ceux observés à 4s et 6s) est observée dans une région postérieure du gyrus temporal supérieur gauche.
- Des décroissances de l'activité neuronale de 4s à 6s ont été observées au sein du cortex auditif primaire, du gyrus temporal supérieur et de la partie postérieure de l'insula, de manière bilatérale, et au sein de la partie triangulaire du gyrus frontal inférieur droit.
- Aucune croissance significative de l'activité neuronale de 4s à 6s n'a été observée.

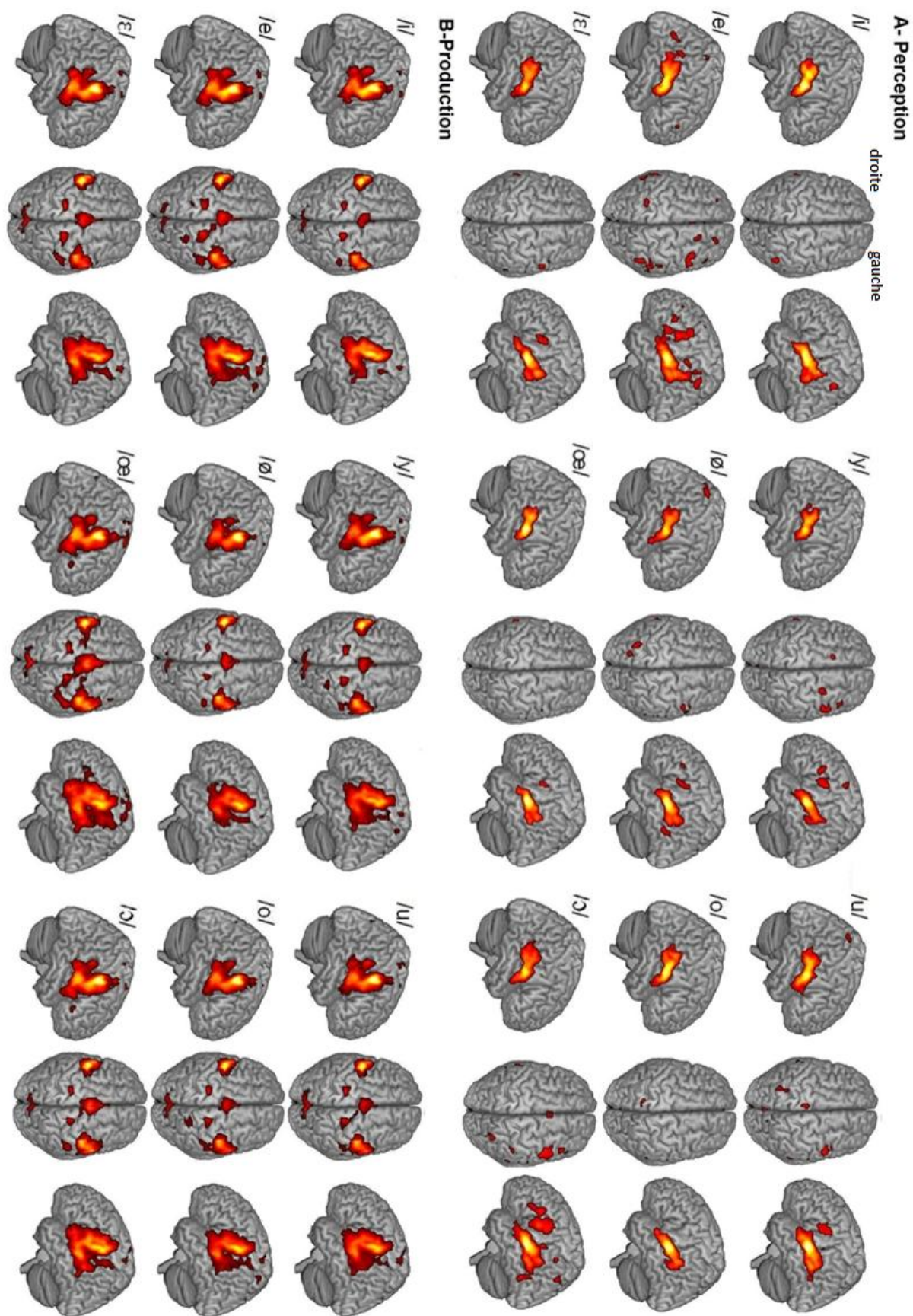


Figure 1.17 : Vue surfacique des régions activées pour les 9 voyelles lors des tâches de perception et de production (FDR, $p < .05$ et taille minimale des clusters de 25 voxels).

Régions	H	BA	Effet principal - Perception				Contrastes estimés			BA	Effet principal - Production				Contrastes estimés		
			x	Y	z	F	4s	5s	6s		x	y	z	F	4s	5s	6s
Régions frontales																	
Aire motrice supplémentaire	L									6	-8	6	54	26.2	0.14	0.03	-0.11
	R									6	8	0	56	20.7	0.24	0.16	0.02
Lobule paracentral	L									6	-16	-18	64	14.7	0.06	0.01	-0.05
Gyrus frontal moyen	L									6	-26	-10	48	15.5	0.12	0.06	-0.07
	R									6	40	-6	52	14.3	0.12	0.08	-0.03
Cortex préfrontal dorsolatéral	R									46	40	16	26	10.9	0.00	-0.10	-0.14
Cortex moteur primaire	R									4	22	-26	66	14.1	0.18	0.14	0.03
Cortex prémoteur	L									6	-40	-6	42	19.6	0.12	0.05	-0.05
	R									6	42	-4	48	15.0	0.09	0.01	-0.07
Prefrontal gyrus	R									9	40	6	28	15.8	-0.04	-0.17	-0.22
Gyrus frontal inférieur	R	45	44	18	18	14.2	0.02	-0.04	-0.06	47	32	28	4	6.9	0.06	0.00	-0.06
Régions temporales																	
Aire temporopolaire	L									38	-36	2	-18	14.5	-0.16	-0.06	0.12
Gyrus temporal supérieur	L	22	-52	-18	6	42.6	0.66	0.57	0.26	22	-58	2	6	19.9	0.05	0.38	0.49
	R	22	52	-16	6	25.1	0.50	0.38	0.18	22	48	8	-2	10.1	0.07	0.20	0.31
Gyrus temporal transverse	L	41	-40	-28	10	32.4	0.37	0.26	0.08	41	-38	-28	10	19.9	0.47	0.42	0.21
	L	42	-56	-30	12	10.2	0.40	0.35	0.19	42	-64	-16	12	19.3	0.17	0.53	0.61
	R	41	52	-34	10	10.2	0.15	0.15	0.07								
Gyrus temporal moyen	R									21	-42	-6	-10	10.8	-0.13	0.07	0.11
Gyrus temporal inférieur	R									37	42	-64	-6	7.2	0.07	0.04	-0.06
Régions pariétales																	
Cortex somatosensoriel	R									3	42	-10	34	10.6	0.57	0.56	0.42
Lobule pariétal inférieur / Gyrus supramarginal	R									40	54	-30	28	13.1	0.15	0.28	0.28
Opércule rolandique	L									43	-62	-6	10	18.1	0.34	0.64	0.72
Lobule pariétal supérieur / Précuneus	L									7	-30	-52	54	18.4	0.12	0.02	-0.11
	R									7	2	-78	30	13.8	0.08	0.32	0.40
Cortex insulaire																	
Insula	L	13	-38	-18	12	17.9	0.19	0.17	0.02								
	R	13	44	-24	16	20.5	0.24	0.16	0.02	13	40	2	2	10.5	-0.05	0.06	0.07
Système limbique																	
Gyrus parahippocampique	L									30	-22	-54	2	18.3	0.05	0.22	0.24
	R									30	12	-42	-4	15.8	-0.10	10.3	0.32
Ganglions de la base																	
Putamen	L										-22	6	12	9.9	0.08	0.09	0.01
Régions occipitales																	
Cortex extra-strié	R									19	34	-64	-8	11.3	0.03	0.04	-0.07

Table 1.8 : Coordonnées dans l'espace MNI des pics d'activation observés pour les tâches de perception et de production montrant des différences d'activités significatives par rapport aux 3 délais d'acquisition (FDR, $p < .05$ et taille minimale des clusters de 25 voxels).

Production

L'effet principal des trois délais d'acquisition montre également pour la tâche de production des décours temporels significativement différents dans certaines régions corticales et sous-corticales (voir Table 1.8 et Figure 1.19).

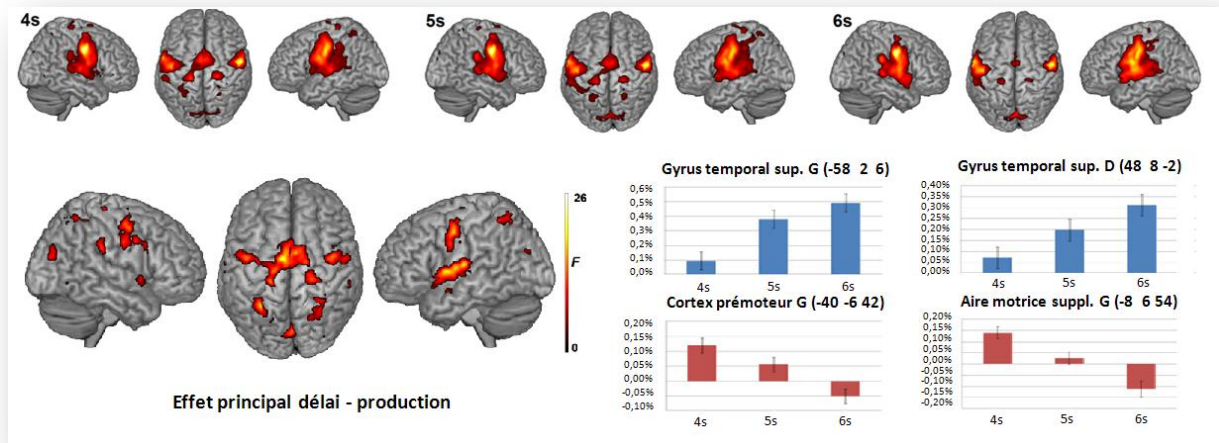


Figure 1.19 : Vue surfacique des régions cérébrales activées lors de la tâche de production selon les 3 délais d'acquisition de 4s, 5s et 6s (haut) et effet principal (bas) montrant les régions cérébrales différemment activées selon les 3 délais d'acquisition. Les diagrammes représentent les valeurs des activations pour les trois délais d'acquisition dans certaines des régions cérébrales indiquées dans l'effet principal ($p < .05$, FDR, et taille minimale des clusters de 25 voxels).

- Des courbes en cloche de l'activité neuronale pour les trois délais d'acquisitions (pic d'activation maximal à 5s et significativement supérieur à ceux observés à 4s et 6s) sont observées dans certaines régions de l'aire motrice primaire ainsi que dans le putamen gauche.
- Des décroissances de l'activité neuronale de 4s à 6s ont été observées au sein des régions frontales incluant l'aire motrice supplémentaire, les cortex prémoteur et moteur primaire et les gyri frontaux moyen et inférieur ainsi qu'au niveau des aires temporelles, visuelles, pariétales supérieures (précuneus/cortex cingulaire postérieur gauche et cortex somatosensoriel primaire droit).
- Des croissances de l'activation neuronale significatives de 4s à 6s ont été observées au niveau des régions temporelles moyennes et supérieures, de l'opercule pariétal gauche et du lobule pariétal supérieur (précuneus) et du cortex cingulaire postérieur.

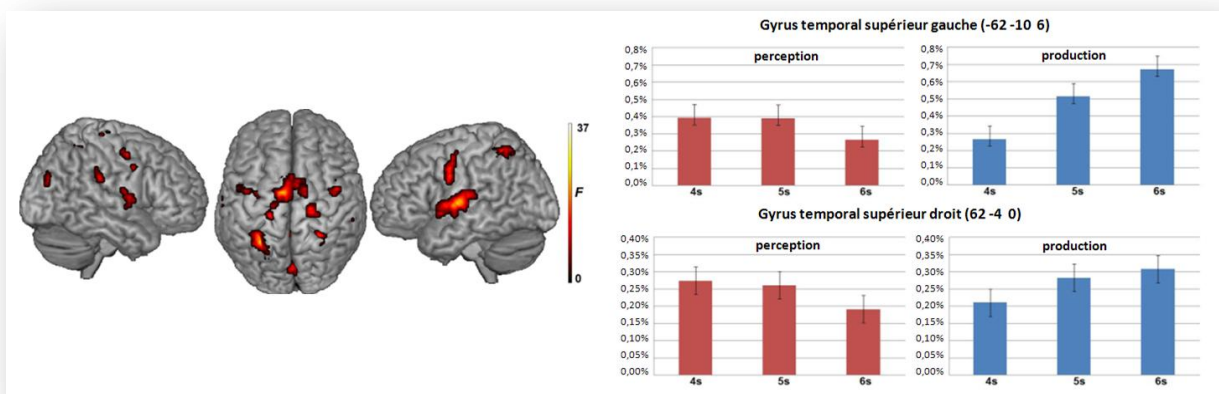


Figure 1.20 : Vue surfacique des régions cérébrales différemment activées lors de la tâche de perception et de production en fonction des délais d'acquisition de 4s, 5s et 6s. Les paramètres estimés des activations pour les trois délais d'acquisition sont indiqués pour certaines régions cérébrales ($p < .05$, FDR, et taille minimale des clusters de 25 voxels).

Interaction Tâche x Délai

Décroissance en production. Une décroissance des activités est observée entre les 3 délais d'acquisition lors de la tâche de production dans les cortex prémoteur et moteur primaire, le gyrus frontal inférieur gauche, le cortex préfrontal droit et, de manière bilatérale, l'aire motrice supplémentaire et le précuneus. En revanche, aucune variation d'activité entre les 3 délais n'y est constatée lors de la tâche de perception (voir Figure 1.20).

Croissance en production. De manière générale, le cortex prémoteur et le gyrus supramarginal droits ainsi que, de manière bilatérale, l'opercule pariétal et le gyrus temporal supérieur montrent des patterns décroissants de 4s à 6s pour la tâche de perception tandis que le signal BOLD augmente pour la tâche de production (voir Figure 1.20).

1.2.4 Discussion

De nombreux modèles neurobiologiques, phonétiques et psycholinguistiques de la perception et/ou de la production de la parole proposent une connexion fonctionnelle entre systèmes sensoriels et moteur (Hickok et Poeppel, 2000, 2004, 2007; Scott et Johnsrude, 2003; Callan et al., 2004; Guenther, 2006; Wilson et Iacoboni, 2006; Skipper et al., 2007; Rauschecker et Scott, 2009; Rauschecker, 2011; Hickok, Houde et Rong, 2011; Guenther et Vladusich, 2012; Perkell, 2012; Schwartz et al., 2012). Lors de la perception de la parole, il est admis que les interprétations phonétiques soient contraintes par des mécanismes de simulation motrice interne basés sur les connaissances articulatoires procédurales de l'auditeur, notamment lors de condition d'écoute défavorable et/ou de séquences complexes de parole coarticulées. Lors de la production de la parole, l'hypothèse est souvent émise que des boucles correctives sensorimotrices contribueraient au repérage d'éventuelles divergences entre les états sensoriels prédits et les retours sensoriels réels dans le but d'assurer un contrôle en ligne de la production. Ces mécanismes de contrôle correctifs seraient fortement recrutés dans le cas d'altération du feedback auditif et/ou lors de la production de séquences de parole, mais ne joueraient pas un rôle crucial lors de la production d'unités de parole simples et/ou surprises (Guenther et Vladusich, 2012).

A la lumière de ces hypothèses, l'objectif de la présente étude était de déterminer l'existence de telles interactions sensorimotrices lors de la perception et production de voyelles maintenues, indépendamment de processus de séquençage syllabique ou de mécanismes de coarticulation entre phonèmes, et en utilisant la technique d'acquisition de sparse sampling afin de limiter l'influence du bruit du scanner sur le traitement auditif. Il s'agissait ainsi de clarifier les réseaux neuronaux sous-jacents à la perception et la production de voyelles et de déterminer un possible couplage fonctionnel des systèmes de perception et production. D'autre part le but était de questionner la nature des représentations cérébrales des unités de parole et de là, de déterminer une possible organisation phonémotopique des voyelles ou types de voyelles. Trois résultats majeurs ont été observés.

- Les deux tâches de perception et de production impliquent des activations aussi bien sensorielles que motrices.
- Les deux systèmes de perception et de production partagent un réseau sensorimoteur commun, correspondant à une voie postéro-dorsale dans l'hémisphère gauche.
- Bien qu'aucune somatotopie n'ait été observée, les résultats montrent une implication accrue de certaines structures motrices et somatosensorielles, particulièrement pour les voyelles mi-ouvertes arrondies.

En résumé, nos résultats montrent qu'en dehors du cortex auditif, la voie postéro-dorsale gauche (incluant la partie operculaire de l'aire de Broca, le cortex prémoteur ventral adjacent et la jonction temporo-pariétale) est une partie influente du traitement sensorimoteur des voyelles. En effet, ces

régions ont été observées dans l'analyse de conjonction comme activées aussi bien lors de la perception que lors de la production des voyelles. Néanmoins, alors que des analyses spécifiques des traits phonétiques confirment de nouveau l'implication de la voie postéro-dorsale gauche, aucune ségrégation topographique n'a pu être observée entre voyelles. Ces résultats suggèrent ainsi fortement que les représentations des voyelles seraient distribuées à travers les régions cérébrales sensorimotrices de la voie dorsale et apportent un soutien supplémentaire pour un couplage fonctionnel entre systèmes de perception et de production.

Système de production des voyelles

Bien que les corrélats neuronaux du contrôle moteur de la parole aient été largement étudiés par le passé (Murphy et al., 1997; Wise et al., 1999; Riecker et al., 2000a,b, 2005, 2008; Brown et al., 2005; Guenther, Ghosh et Tourville, 2006; Bohland et Guenther, 2006; Özdemir, Norton et Schlaug, 2006; Sörös et al., 2006; Terumitsu, et al., 2006; Brown, Ngan et Liotti, 2008; Brown et al., 2009; Chang et al., 2009), les études précédentes se sont cependant principalement concentrées sur des tâches de production de syllabes impliquant un séquençage complexe de phonèmes et/ou de syllabes et des mécanismes de coarticulation. Les études relatives à la production de phonèmes, en l'occurrence de voyelles, concernent uniquement la production d'une seule voyelle (Sörös et al., 2006 : /a/; Terumitsu et al., 2006 : /e/; Brown, Ngan et Liotti, 2008 : /ə/; Grabski et al., 2012 : /i/), se sont focalisées sur le cortex moteur primaire (Terumitsu et al., 2006), rapportent des résultats mêlant voyelles et syllabes (Ghosh, Tourville et Guenther, 2008) et/ou, de manière plus importante, utilisent les voyelles comme tâche contrôle afin de déterminer de manière plus spécifique les substrats neuronaux de la production des syllabes, mots ou phrases (Özdemir et al., 2006; Sörös et al., 2006; Ghosh, Tourville et Guenther, 2008). En utilisant 9 voyelles maintenues distinctes, nos résultats permettent ainsi d'affiner les résultats précédemment observés dans ces études sur les processus cérébraux sous-jacents à la production de voyelles.

Dans notre étude, la production des voyelles implique un réseau largement distribué de régions motrices corticales et sous-corticales, incluant les cortex prémoteur et sensorimoteur primaire (s'étendant ventralement vers la partie operculaire du gyrus frontal inférieur, l'opercule rolandique et l'insula gauche et dorsalement vers l'aire motrice supplémentaire), les ganglions de la base, le noyau rouge et le cervelet. Des activations bilatérales du gyrus/sulcus temporal supérieur (incluant les aires auditives primaire, secondaire et associative s'étendant depuis l'aire temporopolaire jusque la jonction temporo-pariétale) ont également été observées, tout comme l'activation de régions des cortex pariétaux supérieur et inférieur, du système limbique (cortex cingulaire antérieur et thalamus) et du cortex visuel.

Ce réseau de production des voyelles est en accord avec de précédentes études sur le contrôle moteur de la parole et sur l'existence d'un réseau minimal pour la production ouverte de la parole (Bohland et Guenther, 2006; voir Chapitre 1). Ce réseau est composé de régions corticales et sous-corticales impliquées dans la préparation et l'exécution motrice, les traitements auditif/phonologique et dans des boucles sensorimotrices régulatrices (pour des revues, voir Jürgens, 2002; Bohland et Guenther, 2006; Guenther, 2006; Sörös et al., 2006; Riecker et al., 2005; Guenther et Vladusich, 2012). Ces régions cérébrales sont traditionnellement assignées à l'initiation ou la suppression des énoncés verbaux (aire motrice supplémentaire et partie antérieure du cortex cingulaire), la planification des séquences de mouvements (cortex prémoteur ventrolatéral, cortex frontal inférieur et insula antérieure gauche), le feedback auditif et le traitement phonologique (gyrus/sulcus temporal supérieur) et finalement, l'exécution motrice impliquant une coordination musculaire fine des mouvements visés et l'innervation des muscles du conduit vocal (système cortico-bulbaire, ganglions de la base, cervelet, noyau rouge). De plus, des processus de régulation en ligne des commandes motrices lors de la production de voyelles pourraient également être impliqués, notamment lors du contrôle temporel des gestes moteurs (cervelet supérieur et ganglions de la base via des projections thalamo-corticales) et/ou lors de mécanismes correctifs en cas de

divergence entre entrées auditives et somatosensorielles prédites et retours sensoriels réels (gyrus/sulcus temporal supérieur, aires somatosensorielles primaire et associative et cortex pariétal inférieur).

Système de perception des voyelles

En plus du système auditif, la perception de voyelles a induit des activations pariétales dans le gyrus supramarginal, le précuneus, l'opercule rolandique droit et des activations frontales dans la partie operculaire du gyrus frontal inférieur, s'étendant jusqu'au cortex prémoteur ventral adjacent, ainsi que des activations additionnelles au niveau de la partie triangulaire dans l'hémisphère gauche. Des activations frontales supplémentaires ont été également observées dans le cortex dorsolatéral préfrontal gauche, le gyrus frontal supérieur gauche et moyen droit. Enfin, des activations ont également été observées dans les ganglions de la base (incluant le noyau caudé et la substance noire), dans les structures limbiques (incluant le thalamus, le cortex cingulaire et l'hippocampe), le noyau rouge et dans les régions visuelles striées et extrastriées.

Traitement acoustico-phonétique des informations auditives : voie antérieure ou postérieure ?

Comme lors de la tâche de production, la tâche de perception de voyelles a induit de larges activations auditives bilatérales au sein du gyrus/sulcus temporal supérieur (incluant le gyrus de Heschl et s'étendant antérieurement vers l'aire temporopolaire, postérieurement vers le planum temporale et la jonction temporo-pariétale, et ventralement vers le gyrus temporal moyen). Ces activations apparaissent en accord avec de précédentes études IRMf sur la perception des voyelles et, plus généralement, avec certains modèles neurobiologiques récents de la perception de la parole (Scott et Johnsrude, 2003; Callan et al., 2004; Guenther et al., 2004; Obleser et al., 2006; Uppenkamp et al., 2006; Wilson et Iacoboni, 2006; Hickok et Poeppel, 2007; Skipper et al., 2007; Rauschecker et Scott, 2009; Hickok, Houde et Rong, 2011; Rauschecker, 2011). Lors de la perception auditive de la parole, il est généralement admis que le signal acoustique de la parole soit traité par des voies parallèles spécialisées dans différents niveaux de traitements. Des premiers traitements acoustiques de bas niveau, non spécifiques aux sons de parole, auraient ainsi lieu dans les gyri temporaux supérieurs dorsaux bilatéraux (incluant le gyrus de Heschl et des régions du planum temporale). La localisation exacte des régions cérébrales spécialisées dans le décodage acoustico-phonétique et la catégorisation des phonèmes perçus reste néanmoins controversée (pour une revue récente, voir Turkeltaub et Coslett, 2010). En effet, certains chercheurs proposent que les signaux acoustiques complexes (non spécifiques à la parole) sont d'abord traités dans la partie médio-latérale du gyrus/sulcus temporal supérieur et que les représentations phonétiques des sons de parole seraient instanciées suite à ces traitements dans la partie antérieure du gyrus/sulcus temporal supérieur (Scott et Johnsrude, 2003; Obleser et al., 2006; Rauschecker et Scott, 2009; Rauschecker, 2011). Au contraire, d'autres chercheurs soutiennent que les catégories phonétiques sont instanciées dans la partie postérieure du gyrus temporal moyen et au sein de la partie adjacente du sulcus temporal supérieur (Hickok et Poeppel, 2007; voir aussi Belin et Zatorre, 2000). En dehors de ces modèles, des études suggèrent également des localisations cérébrales différenciées pour les traitements phonétiques mis en œuvre lors de la perception de différentes voyelles. Par exemple, Obleser et collègues (2006) ont observé que la perception de séquences de voyelles provoque des activations accrues de la partie antérieure du gyrus temporal supérieur par rapport à l'écoute de séquences non langagières ainsi qu'une distinction topographique dans cette région entre représentations de voyelles antérieures et postérieures³⁹. Guenther et collègues (2004) ont au contraire observé des activités accrues dans les gyri temporaux moyens antérieur et postérieur lors de la perception d'exemples prototypiques d'une catégorie de voyelle (/i/) par rapport à des voyelles non

³⁹ Il est cependant à noter que ce résultat était issu d'une analyse à effets fixes et donc non généralisable.

prototypiques n'étant pas clairement appariées à une catégorie phonologique. Enfin, en contrastant la perception de voyelles avec des stimuli non langagiers, Upperkamp et collègues (2006) ont observé des activations bilatérales spécifiques aux voyelles au sein du sulcus temporal supérieur, le long du gyrus temporal supérieur dans la direction antéropostérieure et inférieure par rapport au gyrus de Heschl. Comme l'ensemble des régions mentionnées dans ces études et modèles ont été trouvés activées dans l'étude présente, nos résultats ne permettent ni d'appuyer ni de contredire l'existence d'une voie auditive antérieure (Scott et Johnsrude, 2003; Rauschecker et Scott, 2009; Rauschecker, 2011) ou postérieure (Hickok et Poeppel, 2007) impliquée dans les processus de décodage acoustico-phonétique lors de la perception de voyelles.

Implication de la voie dorsale

D'autres activations ont été observées, principalement dans l'hémisphère gauche, dans la partie operculaire du gyrus frontal inférieur, s'étendant au cortex prémoteur ventral adjacent, dans la partie triangulaire du même gyrus frontal inférieur, ainsi que dans la jonction temporo-pariétale (incluant la partie postérieure du gyrus/sulcus temporal supérieur et la partie ventrale du gyrus supramarginal⁴⁰). Ces activations apparaissent en accord avec l'existence d'une voie de traitement postéro-dorsale dans l'hémisphère gauche lors de la perception de parole (Hickok et Poeppel, 2000, 2004, 2007; Scott et Johnsrude, 2003; Callan et al., 2004; Wilson et Iacoboni, 2006; Skipper et al., 2007; Rauschecker et Scott, 2009; Rauschecker, 2011) liant les représentations auditives avec les représentations articulatoires de la parole dans la partie postérieure du gyrus frontal inférieur et le cortex prémoteur ventral via le lobule pariétal inférieur (Rauschecker et Scott, 2009; Rauschecker, 2011) ou via l'aire 'SPT' (pour 'sylvian-parieto-temporal', localisée à la jonction de la frontière pariéto-temporale; Hickok et Poeppel, 2007). Ces modèles postulent en effet l'existence, par des projections sensori-motrices et motori-sensorielles successives, d'un mécanisme de simulation motrice dont le but est de prédire les conséquences sensorielles des commandes motrices activées en fonction des représentations auditives et de contraindre l'interprétation phonétique finale de l'auditeur. Ainsi, Skipper et collègues (2007) proposent que les entrées sensorielles de la parole (auditives et/ou visuelles) induisent l'activation initiale de candidats phonétiques possibles contenus dans le signal de parole dans le gyrus temporal supérieur gauche. Ces représentations phonétiques seraient alors appariées aux commandes motrices correspondantes par un mécanisme inférentiel sensorimoteur interne localisé dans la partie postérieure du gyrus frontal inférieur gauche. Les informations relatives à ces commandes motrices seraient ensuite envoyées au cortex prémoteur ventral et au cortex sensorimoteur primaire en vue de générer, en retour, des prédictions sur les conséquences acoustiques et somatosensorielles relatives à l'exécution simulée de ces commandes par des mécanismes de copies d'efférence. Enfin, ces prédictions seraient renvoyées à la partie postérieure du gyrus temporal supérieur gauche, où elles seraient comparées avec les hypothèses phonétiques initiales pour aboutir à une décision phonétique finale quant aux phonèmes perçus. De façon similaire, Rauschecker et Scott (2009; voir aussi Rauschecker, 2011) proposent que suite à des traitements acoustique de bas niveau dans le gyrus de Heschl, la perception de la parole implique des processus de décodage acoustico-phonétique dans la voie antérieure ventrale allant jusqu'au cortex temporal antérieur et, de là, aux régions frontales inférieures supposées contenir des

⁴⁰ Dans une méta-analyse récente, Obleser et Eisner (2009) ont reporté cinq papiers montrant des activations du gyrus supramarginal reliées à des traitements phonologiques et proposent que cette région constitue l'unité de stockage phonologique ('phonological store') en mémoire de travail verbale. Une autre méta-analyse ALE (Turkeltaub et Coslett, 2010) discute des diverses possibilités quant au rôle fonctionnel du gyrus supramarginal en perception de la parole. En dehors de la mémoire de travail, le gyrus supramarginal pourrait avoir un rôle plus général dans la catégorisation/discrimination phonétique, de concert avec le gyrus angulaire, et pourrait fonctionner en tant que site amodal intégratif de plus haut niveau ('amodal integrative processing hub') par rapport aux entrées multisensorielles convergentes. La troisième possibilité, proposée dans le modèle DIVA (Guenther, 2006; Guenther et Vladusich, 2012), est que le gyrus supramarginal jouerait un rôle dans le contrôle en ligne de la production de la parole, en comparant les copies d'efférences prédictives somatosensorielles émises par les régions motrices aux retours sensoriel réel. Le même mécanisme pourrait être exploité en perception de la parole pour contraindre les interprétations phonétiques des signaux de parole perçus ambigus (voir Skipper et al., 2007; Rauschecker et Scott, 2009)

représentations invariantes des catégories phonétiques. Ces informations seraient ensuite envoyées au cortex prémoteur ventral pour être appariées avec les représentations articulatoires correspondantes. Les conséquences sensorielles de l'exécution simulée de telles commandes motrices seraient alors envoyées au lobule pariétal inférieur par le biais de copies d'efférence. En direction inverse, une voie dorsale impliquerait le lobule pariétal inférieur en tant que région d'intégration sensorimotrice, où des comparaisons entre les copies d'efférence prédictives provenant du cortex prémoteur ventral et les informations sensorielles réelles auraient lieu et permettraient une désambiguïsation de l'information phonétique⁴¹.

Autres activations observées

- **Ganglions de la base (striatum) :** Contrairement aux résultats reportés dans la plupart des études trouvés dans la littérature sur la perception de la parole, nous avons également observé des activations du striatum et de la substance noire. Il a été montré que le circuit dopaminergique (incluant les ganglions de la base, la substance noire et le thalamus, régions activées dans la présente étude) est impliqué dans l'apprentissage sensorimoteur (Doya, 1999) et, plus particulièrement pour le striatum, dans l'apprentissage perceptif catégoriel (Seger et al., 2005, 2006). Face à cette possibilité, Callan et al. (2003) interprètent les activations des ganglions de la base lors de la perception de la parole comme reliées à des mécanismes d'apprentissage permettant une mise à jour et une sélection des représentations sensorielles et motrices des unités de parole perçues. Bien qu'une tâche d'écoute passive ait été utilisée, il n'est pas exclu que les participants aient volontairement essayé d'identifier les voyelles perçues. De plus, comme les voyelles mi-ouvertes ne sont habituellement pas perçues en contexte isolé, celles-ci auraient pu entraîner un effort d'apprentissage supplémentaire dans ces processus de discrimination. Une explication alternative concerne l'implication possible du striatum dans la gestion des ressources attentionnelles et de la mémoire de travail (verbale) phonologique (Teichmann et al., 2009) lors de la perception de la parole⁴².
- **Activations somatosensorielles :** Bien que n'apparaissant pas de manière robuste lors de la perception des voyelles, des activations au sein du cortex somatosensoriel primaire sont néanmoins observées. Ce résultat, bien qu'à prendre avec précaution, est en accord avec différentes études montrant une activation du cortex somatosensoriel lors de la perception de parole, notamment visuelle ou audio-visuelle (par exemple, Möttönen et al., 2004; Skipper et al., 2007).
- **Activations visuelles :** De manière intéressante, nous observons également des activations occipito-temporales visuelles au sein du gyrus fusiforme lors de la perception des voyelles. Une hypothèse possible pourrait être celle de liens associatifs entre codes orthographiques et représentations phonologiques des voyelles, pouvant notamment découler des étapes d'acquisition de la lecture et de l'écriture (où une production verbale concomitante est souvent présente). Ces liens associatifs pourraient ainsi expliquer le recrutement en ligne et de manière irrépressible des représentations orthographiques correspondantes aux voyelles perçues (Raij et

⁴¹ Une caractéristique importante de notre étude concerne l'écoute de voyelles propres aux sujets. Bien que de nombreuses études se soient intéressées à la reconnaissance visuelle de soi ou de ses propres actions (par exemple, Decety et Sommerville, 2003; Sugiura et al., 2005; Uddin et al., 2006; Kaplan et al., 2008; Hodzic et al., 2009), des études récentes sur la perception de notre propre voix (par exemple, Nakamura et al., 2001; Kaplan et al., 2008; Rosa et al., 2008) ont également montré des activations dans l'hémisphère droit au niveau des sites frontaux et parainsulaires. La plupart de ces études considèrent le gyrus frontal inférieur droit comme un site amodal sous-tendant la conscience de soi. Ainsi Uddin et al. (2005, 2006) soutiennent que les sites frontal et pariétal inférieurs sous-tendent le sens de l'agentivité et les processus de discrimination de soi vs l'autre, et Decety et Grèzes (2006) ont révélé dans une méta-analyse de 14 études portant sur la théorie de l'esprit et de 11 études portant sur l'agentivité une implication forte du lobule pariétal inférieur, proche de la frontière temporelle.

⁴² Un rôle fonctionnel des ganglions de la base lors de la perception de la parole a été proposé récemment par Kotz et al. (2009) et concerne leurs implications dans la synchronisation des aspects temporels et séquentiels des sons perçus. Néanmoins les auteurs discutent ces activations dans le contexte de traitements syntaxiques, faiblement relié à la perception phonémique.

al., 1999; Peereman et al., 2009; Dehaene et al. 2010). De manière encore plus surprenante, nous avons également observé des activations au sein du cortex visuel strié et extra-strié. Ces activations pourraient correspondre à des processus d'imagerie visuelle orthographique évoqués par les phonèmes perçus. En appui de cette hypothèse, Peereman et al. (2009) soutiennent l'idée que les connaissances orthographiques, acquises lors de l'apprentissage de la lecture et de l'écriture, puissent influencer la perception auditive de mots. Bien que leurs travaux concernent plus les processus de catégorisation sémantique, ils appuient également l'idée selon laquelle les connaissances orthographiques puissent contraindre les structures phonologiques correspondantes. De là, il est possible que lors de la perception auditive de voyelles, une évocation orthographique visuelle irrépressible puisse avoir lieu.

Activations différenciées lors de la perception et la production des voyelles

La comparaison des activations observées lors de la production et perception des voyelles montre 3 types de différences des réponses hémodynamiques.

- Premièrement, des activations significatives lors de la tâche de production et quasi nulles lors de la tâche de perception sont observées dans les cortex moteur primaire et prémoteur, la partie caudale de l'aire motrice supplémentaire (SMA-proper), le gyrus supramarginal droit, le cortex insulaire, les ganglions de la base (noyau caudé et globus pallidus), certaines régions du système limbique (cortex cingulaire antérieur et thalamus) et dans le cortex visuel.
- Deuxièmement, certaines régions non activées lors de l'écoute des voyelles apparaissent déactivées lors de la tâche de production de voyelles par rapport à la condition de base (cortex dorsolatéral préfrontal et supérieur frontal, gyri frontaux moyen et médian, partie triangulaire du gyrus frontal inférieur, gyrus angulaire et gyrus supramarginal adjacent, partie postérieure des gyri temporal moyen et supérieur, lobule pariétal supérieur et cortex cingulaire postérieur). Ces régions cérébrales, montrant un niveau d'activité moindre lors de la tâche de production par rapport à la condition de repos, sont traditionnellement considérées comme faisant partie d'un réseau 'par défaut' ('default mode network'; Raichle et al., 2001). Il est en effet admis qu'une condition de repos implique une activité cérébrale organisée incluant des processus verbaux, d'imagerie visuelle et de conscience de soi (Binder et al., 1999; Gusnard et al., 2001; Mazoyer et al., 2001; Raichle et al., 2001, McKiernan et al., 2003). Bien que cette question soit débattue, certains auteurs proposent que les déactivations induites par une tâche ('task-induced deactivations') résultent de mécanismes de réallocations des ressources attentionnelles, permettant un recrutement optimal d'autres régions dans la réalisation de la tâche (McKiernan et al., 2003, 2006).
- Enfin, certaines régions temporales et pariétales (gyri temporaux supérieur et transverse, opercule rolandique adjacent) montrent une activité accrue par rapport à la ligne de base lors des tâches de perception et de production, mais de plus amples activations lors de la tâche de production. Ce résultat apparaît en conflit avec des études IRMf précédentes montrant une activité moindre au niveau du cortex auditif lors de la production ouverte de séquences de parole par rapport à l'écoute de ces mêmes séquences (Christoffels, Formisano et Schiller, 2007; Christoffels et al., 2011). Bien qu'une procédure d'appariement du volume ait été utilisée dans notre étude, afin d'apparier le volume des stimuli auditifs dans la tâche de perception et celui du feedback auditif dans la tâche de production, une explication plausible à ce résultat est celle de différences auditives dues à la conduction osseuse et aux contractions de l'oreille moyenne.

Couplage fonctionnel des systèmes de perception et de production des voyelles

Comme confirmé dans l'analyse de conjonction, des activations sensorimotrices communes lors de la perception et la production de voyelles ont été observées au sein du gyrus/sulcus temporal supérieur, s'étendant dorsalement dans l'hémisphère gauche à la jonction temporo-pariétale, de la partie operculaire du gyrus frontal inférieur gauche et du cortex prémoteur ventral adjacent. Ces résultats appuient ainsi l'hypothèse d'une connexion fonctionnelle entre systèmes sensoriel et moteur aussi bien lors de la perception que lors de la production de voyelles. Il est intéressant de noter que ces activités sensorimotrices sont observées dans des conditions d'écoute favorables; c'est-à-dire impliquant des unités élémentaires de la parole peu contaminées par des mécanismes de coarticulation, propres à chaque locuteur et n'impliquant donc pas ou peu de processus de normalisation dans l'espace acoustique de l'auditeur, et perçues en dehors du bruit du scanner. Ce résultat suggère par conséquent une implication quasi-systématique du système moteur lors de la perception de parole.

De manière importante, bien que notre étude démontre clairement le recrutement du système moteur lors de la perception passive de voyelles, ces résultats sont intrinsèquement corrélationnels et ne peuvent donc pas être utilisés comme preuve d'un rôle fonctionnel médiateur du système moteur dans la compréhension des voyelles. Sur ce point, des études par stimulations électrocorticales lors d'opérations neurochirurgicales ou par TMS et les données cliniques issues de l'étude de patients aphasiques ne permettent pas de conclure véritablement à un rôle causal fonctionnel du système moteur dans le traitement auditif de la parole dans des conditions optimales d'écoute (pour une revue, voir Sato, Tremblay et Gracco, 2009). En effet, bien que des études TMS récentes ont montré qu'une perturbation temporaire de l'activité du système moteur modifie l'aptitude des participants à catégoriser des syllabes ambiguës ou présentées dans du bruit (Meister et al., 2007; d'Ausilio et al., 2009; Möttönen et Watkins, 2009), de tels effets d'interférence n'ont pu être confirmés dans des tâches de discrimination de syllabes clairement perçues (Sato, Tremblay et Gracco, 2009; d'Ausilio et al., 2011). Il est de plus à noter que malgré l'activation commune de la partie postérieure du gyrus frontal inférieur et du cortex prémoteur ventral adjacent lors de la perception et production des voyelles, des différences d'activité ont été observées dans ces régions entre voyelles mi-ouvertes, d'une part, et voyelles fermées et mi-fermées, d'autre part lors de la tâche de production mais non lors de celle de perception. Ce résultat peut être relié à une étude IRMf récente de perception et de production de la parole (Tremblay et Small, 2011), où une modulation de l'activité neuronale due à la complexité articulatoire a été observée en production mais non en perception. Pour les auteurs, ce résultat ne confirme pas le caractère spécifique des activations motrices observées lors de la perception de la parole.

A partir de ces résultats, étant donné le fait que des activités motrices ont été observées lors de la perception des voyelles indépendamment de processus de séquençage syllabique, de mécanismes de coarticulation entre phonèmes et une influence limitée du bruit du scanner sur le traitement auditif, il devient nécessaire de discuter de possibles fonctions et rôles alternatifs du système moteur lors de la perception de parole dans ces conditions favorable d'écoute. Comme mentionné précédemment, même les opposants à tout rôle causal du système moteur dans les traitements de décodage acoustico/phonétique et de compréhension de la parole ne remettent pas en question l'existence d'interactions entre systèmes de perception et de production de la parole (voir Diehl, Lotto et Holt, 2004; Lotto, Hickok et Holt, 2009; Hickok, Houde et Rong, 2011). En effet, les activités motrices observées lors de la perception de la parole pourraient s'expliquer du fait de la coactivation des régions auditives et motrices lors de la production de la parole, coactivation pouvant mener à la formation de liens spécifiques articulo-auditifs par le biais d'apprentissage associatif et de principes hebbiens (Pulvermüller et al., 2006; Lotto, Hickok et Holt, 2009) ce, notamment lors de l'acquisition de la parole pendant l'enfance. De ce point de vue, l'activation de la voie dorsale observée dans la

présente étude lors de la perception de voyelles n'impliquerait pas forcément sa participation dans des traitements phonétiques mais représenterait plutôt la conséquence de processus d'apprentissage associatif antérieur.

Une autre possibilité est que l'activité motrice observée lors de la perception de voyelles pourrait signifier un cas spécifique de convergence sensorimotrice automatique (Scott, McGettigan et Eisner, 2009; Sato, Tremblay et Gracco, 2009). Chez les adultes, il a été démontré, que des effets de convergence ont lieu systématiquement lors d'interactions conversationnelles et qu'ils se manifestent sous diverses formes, par exemple par la posture (Shockley et al., 2003), les mouvements de tête et les expressions faciales (Estow et al., 2007), les mouvements respiratoires (McFarland, 2001), l'intensité vocale (Natale, 1975) et le débit de parole (Giles et al., 1991). Au niveau phonétique, de précédentes études ont également souligné pour l'auditeur une tendance à imiter un certain nombre de caractéristiques phonétiques issues des signaux de parole du locuteur – c'est le phénomène de convergence phonétique. Alors que certaines études ont démontré l'existence d'un tel mécanisme de convergence phonétique dans un cadre écologique et naturel, lors d'interactions conversationnelles (par exemple, Sancier et Fowler, 1997; Pardo, 2006; Aubanel et Nguyen, 2010), cet effet apparaît également dans le cadre d'études en laboratoire, sans interlocuteur ni interactions conversationnelles mais par la simple exposition à des stimuli de parole auditifs et/ou visuels (par exemple, Kerzel et Bekkering, 2000; Gentilucci et Cattaneo, 2005; Gentilucci et Bernardis, 2007; Delvaux et Socquet, 2007; Kappes et al., 2009). Dans ce cadre, une hypothèse quelque peu spéculative serait que l'activité motrice observée lors de la perception de voyelles pourrait représenter des processus sensorimoteurs automatiques d'adaptation aux stimuli perçus. De là, cette hypothèse pourrait indiquer un rôle central du système moteur lors de conversations interactionnelles, un aspect essentiel du langage humain dans son emploi le plus élémentaire ('for motor representations and processes in conversation, an essential aspect of human language in its most basic use' Scott, McGettigan et Eisner, 2009).

Représentations différenciées entre voyelles

Perception : Si les analyses spécifiques aux traits phonétiques confirment l'implication de la voie dorsale dans le traitement auditif des voyelles, elles n'ont révélé en revanche aucune ségrégation en terme topographique entre les différents traits. Ce résultat nul apparaît cohérent avec une récente étude IRMf (Obleser et al., 2010) dans laquelle des analyses univariées, basées sur une soustraction des activations observées lors d'une tâche de perception de parole par rapport à celles observées dans une condition de repos, n'ont pas apporté d'éléments en faveur de différences d'activation entre catégories phonétiques (mais voir Obleser et al., 2006 pour une étude semblable utilisant une analyse à effet fixe). En revanche, en utilisant une méthode d'analyse multivariée de reconnaissance statistique des patrons d'activité sur les mêmes données, Obleser et collègues (2010) ont démontré une classification supérieure au hasard des catégories vocaliques et consonantiques dans le cortex auditif temporal supérieur. Ce résultat apparaît en adéquation avec l'étude princeps effectuée par Formisano et collègues (2008) utilisant également une méthode d'analyse multivariée et démontrant des cartes de discrimination pour trois différentes voyelles isolées (/i/, /a/, /u/) au niveau des aires temporales supérieures, depuis le gyrus latéral de Heschl (parties antérieure et postérieure) jusqu'au sulcus temporal supérieur. Face à ces études, nos résultats soulignent ainsi les limites actuelles d'une approche localisationniste basée sur des méthodes d'analyses univariées et plaident indirectement en faveur de représentations plus largement distribuées au travers de régions cérébrales sensorielles et motrices.

Production : Les analyses spécifiques aux traits phonétiques d'ouverture (voyelles fermées, mi-fermées et mi-ouvertes) et d'arrondissement/antériorité (voyelles antérieures non arrondies, arrondies et postérieures arrondies) confirment l'implication de régions cérébrales communes aux différents traits mais ne révèlent aucune ségrégation topographique entre eux. Néanmoins, alors qu'aucune région cérébrale ne montre de changements significatifs d'activité pour les dimensions

d'arrondissement/antériorité, des activations accrues ont été observées au niveau de la partie postérieure du gyrus frontal inférieur, de l'insula antérieure gauche, du cortex prémoteur bilatéral, de l'aire motrice supplémentaire, du gyrus supramarginal et du cortex somatosensoriel pour les voyelles mi-ouvertes, en comparaison avec les voyelles fermées et mi-fermées. À partir de ce résultat, il est à noter qu'en Français, par rapport aux autres voyelles, les voyelles mi-ouvertes (notamment le /œ/ et le /ɔ/) sont rarement produites en contexte isolé, mais apparaissent plutôt en contexte syllabique fermé (Durand, Laks et Lyche, 2002). Il est possible que cette moindre naturalité de production dans un contexte isolé ait pu influencer la précision motrice de leur réalisation. La distribution et variabilité plus large des valeurs du F1 pour les voyelles mi-ouvertes en comparaison aux voyelles fermées et mi-fermées vont d'ailleurs dans le sens de cette proposition. De manière intéressante, cette augmentation d'activité a été observée non seulement dans les régions cérébrales classiquement dédiées à l'initiation et la planification motrices (aire motrice supplémentaire, gyrus frontal inférieur et cortex prémoteur, cortex insulaire antérieur) mais aussi dans le cortex somatosensoriel secondaire et le lobule pariétal inférieur. Dans le modèle neurocomputationnel DIVA de production de la parole (Guenther, 2006; Guenther et Vladusich, 2012), il est supposé que des cartes d'états et d'erreurs somatosensorielles, localisées dans ces régions pariétales, comparent les buts somatosensoriels prédits (soit les sensations tactiles et proprioceptives attendues des unités de parole produites) avec le feedback somatosensoriel réel du conduit vocal. Les sorties de ces cartes d'erreurs somatosensorielles se propagent ensuite vers le cortex moteur dans le but de permettre, si nécessaire, une correction en ligne des commandes motrices du locuteur. En accord avec cette hypothèse, Golfinopoulos et collègues (2011) ont observé une augmentation d'activation dans le gyrus supramarginal antérieur de manière bilatérale lorsque le feedback somatosensoriel du sujet lors d'une tâche de production de parole était modifié de manière soudaine ou inattendue par des perturbations articulatoires de la mâchoire. Dans notre étude, l'activation accrue du cortex somatosensoriel et du lobule pariétal inférieur lors de la production de voyelles mi-ouvertes pourrait donc refléter une demande accrue des processus de monitoring somatosensoriel dans le but de mieux contrôler la production de ces voyelles. De manière intéressante, de telles modulations d'activation n'ont pas été observées au niveau du cortex auditif, ce qui suggère des boucles de feedback corrective différentes pour les processus de régulation et surveillance somatosensorielle ou auditive. Il est important de noter que des activations moins fortes ont été observées dans ces régions pour les voyelles fermées et mi-fermées et ce non seulement par comparaison aux voyelles mi-ouvertes mais aussi par rapport à la condition de repos. Ce résultat suggère une diminution d'activité de ces régions pariétales lors de la production d'unités de parole surappries. Face à cette hypothèse, ces résultats appuieraient l'hypothèse de l'existence de mécanismes inhibiteurs moteurs vers ces régions, correspondant à une réduction d'activité au niveau du lobule pariétal inférieur et du cortex somatosensoriel dans le cas de production de voyelles surappries mais également une augmentation de l'activité de ces régions dans des circonstances moins naturelles nécessitant des processus de monitoring plus importants des actions orofaciales produites (Guenther, 2006; Guenther et Vladusich, 2012; Golfinopoulos et al., 2011; voir aussi Christoffels et al., 2011).

Délais d'acquisition

L'analyse de groupe relative aux 3 délais d'acquisition des images fonctionnelles révèle des activations largement communes aux 3 délais d'acquisitions (4s, 5s, 6s) lors des tâches de perception et de production de la parole, confirmant par là même la bonne adéquation de ces délais dans l'acquisition du signal BOLD lors de ces deux tâches. Néanmoins, l'effet principal des trois délais d'acquisition montre également des décours temporels significativement différents dans certaines régions corticales et sous-corticales et ce, également, en fonction de la tâche. Quatre patterns de modification du signal BOLD en fonction de la tâche ont ainsi émergé :

- *Perception – activité croissante de 4s à 6s, production – activité décroissante de 4s à 6s.* Cela est

les cas pour les régions pariétales supérieures, le claustrum et le putamen droit.

- *Perception – pas de différence, production – activité croissante de 4s à 6s.* Ce pattern d'activité est observé au niveau des aires occipitales et peut être expliqué par la stimulation visuelle continue de la consigne visuelle dans la tâche de production, absente dans la tâche de perception.
- *Perception – pas de différence, production – activité décroissante de 4s à 6s.* Cela est le cas pour les régions frontales, y compris l'aire motrice supplémentaire, le cortex frontal inférieur, moyen, prémoteur et moteur primaire. En production, ces régions sont impliquées dans la boucle préparatrice pour planifier les séquences motrices (cortex prémoteur) et initier (aire motrice supplémentaire) l'exécution du mouvement (cortex moteur primaire).
- *Perception – activité décroissante de 4s à 6s, production – activité croissante de 4s à 6s.* Cela est le cas pour le gyrus temporal supérieur, les régions pariétales inférieures incluant le gyrus supramarginal, l'opercule rolandique, l'insula, le thalamus et le cervelet. Ce dernier serait nécessaire, dans la tâche de production de voyelles, dans la surveillance en ligne de l'exécution motrice des voyelles maintenues en fonction des retours sensoriels, auditifs et somatosensoriels, parvenant depuis les régions temporales et pariétales via le thalamus. Un résultat intéressant clairement dépendant de la tâche concerne le gyrus temporal supérieur dans sa partie antérieure proche de l'aire temporopolaire, où le signal BOLD décroît en perception mais augmente en production en fonction du délai.

A l'instar de quelques études qui montrent des différences d'activité de certaines régions cérébrales en fonction de la latence d'acquisition (Riecker et al., 2005; Dresel et al., 2005; Özdemir, Norton et Schlaug, 2006; voir aussi l'étude 1.1), ces résultats apportent des informations importantes quant aux délais temporels des réponses hémodynamiques des diverses régions cérébrales en fonction des tâches effectuées. Ils peuvent donc être intéressants du point de vue méthodologique pour de futures études IRMf utilisant le protocole d'acquisition sparse sampling, afin de capturer les réponses hémodynamiques maximales en fonction des tâches et des régions d'intérêt. Par exemple, nos résultats concordent avec ceux de Gracco, Tremblay et Pike (2005) et de l'étude 1.1 où les réponses hémodynamiques optimales pour la plupart des régions cérébrales concernent un délai d'acquisition de 5 ou 6 secondes. Pour les tâches de perception de parole, les délais de 4 ou 5 secondes apparaissent optimaux⁴³.

Face à des différences d'activation en fonction du délai d'acquisition et de la tâche réalisée, il serait tentant d'émettre des hypothèses fonctionnelles entre le déroulement temporel du signal BOLD et la connectivité fonctionnelle des régions activées (voir la discussion du premier chapitre), nous rappelons cependant que ce genre d'inférence doit être considéré avec précaution puisque la relation entre la décharge neuronale et l'activité BOLD n'est pas linéaire et que d'autres facteurs (métaboliques, physiologiques, etc...) entrent en compte dans les interactions entre ces processus (voir Vanzetta et Grinvald, 2008, ainsi que la discussion du premier chapitre).

1.2.5 Conclusion

La présente étude a permis de mettre en évidence les substrats neuronaux sous-jacents à la production et la perception de (ses propres) voyelles. En accord avec nos hypothèses sur une nature sensorimotrice de ces unités de parole, outre des activations du cortex auditif lors de ces deux tâches, nous avons également observé des activations frontales motrices communes. L'analyse de conjonction démontre ainsi un réseau sensorimoteur commun aux tâches de perception et de

⁴³ Il est important de rappeler que nous avons cherché à apparier au plus près les latences d'acquisition dans les deux tâches. Pour ce faire, dans la tâche de production de voyelles, nous avons pris en compte un délai supplémentaire de 500ms afin de tenir compte du délai entre l'affichage de la consigne visuelle et le début de la production.

production incluant le gyrus frontal inférieur, le gyrus temporal supérieur s'étendant de la région antérieure jusqu'à la frontière temporo-pariétal et la partie ventro-antérieure du gyrus supramarginal. Enfin, bien que nous n'ayons pu déterminer d'organisation phonémotopique des voyelles par l'utilisation d'analyses univariées, un projet en cours utilisant des analyses multi-variées de patterns multi-voxels devrait permettre d'affiner cette question. Par ailleurs, nos résultats ont montré une activité accrue de régions impliquées dans le contrôle moteur de la parole (la partie postérieure du gyrus frontal inférieur, de l'insula antérieure gauche, du cortex prémoteur bilatéral, de l'aire motrice supplémentaire, du gyrus supramarginal et du cortex somatosensoriel) lors de la production des voyelles mi-ouvertes, en fonction du degré de maîtrise articulatoire. Pris ensemble, ces résultats suggèrent un couplage fonctionnel entre systèmes de perception et de production de la parole et appuient l'idée d'une nature sensorimotrice des représentations cérébrales des unités de la parole.

- 2^{nde} Partie -

Contrôle sensorimoteur adaptatif
et codage du but des actions
orofaciales et vocaliques

Avant-propos

Les études présentées dans la première partie de cette thèse ont permis de mettre en évidence les réseaux neuronaux sous-jacents à la réalisation de gestes orofaciaux ainsi qu'à la production et à la perception des voyelles du Français.

Les études 2.1 et 2.2 présentées dans la seconde partie de cette thèse impliquaient des protocoles expérimentaux et des méthodes d'acquisition et de prétraitements des données fonctionnelles et anatomiques en tous points identiques, à l'exception de l'utilisation d'un paradigme d'adaptation. Ce paradigme d'adaptation, dit de répétition suppression, consiste en une diminution de l'activité de populations neuronales spécifiques lors de la production et/ou de l'écoute répétée de stimuli. Dans le cadre de possibles mécanismes adaptatifs reliés aux informations sensorielles et motrices perçues et/ou produites, nous avons utilisé ce paradigme de répétition suppression de manière à déterminer les régions possiblement impliquées dans le codage du but des actions perçues ou produites et dans des processus sensoriels prédictifs/adaptatifs.

Ces deux études ont été réalisées en collaboration avec Marc Sato (Département Parole & Cognition, Laboratoire GIPSA-lab, UMR CNRS 5216 & Université de Grenoble) et Laurent Lamalle (Centre Hospitalier Universitaire de Grenoble, Structure Fédérative de Recherche N°1 "RMN Biomédicale et Neurosciences", Unité IRM 3T).

Ces études ont été principalement subventionnées par un financement de l'Institut National Polytechnique de Grenoble (Bonus Qualité Recherche "Modyc", "Modélisation dynamique de l'activité cérébrale") et par des financements de l'Agence Nationale de la Recherche (ANR "SPIM", "L'imitation dans la parole : des intégrations sensorimotrices à la dynamique des interactions conversationnelles") et par Northern Digital Inc. ("Best student paper award", Ninth International Seminar on Speech Production, Montréal, Canada).

Publications pertinentes

- Grabski K, Lamalle L & Sato M. (sous presse). Somatosensory-motor adaptation of orofacial actions in left posterior parietal and ventral premotor cortices. *Plos One*.
- Grabski K, Lamalle L & Sato M. (en préparation). Mirror mechanisms and sensorimotor adaptive changes during vowel perception and production.

Chapitre 2.1

Adaptation somatosensorimotrice et codage du but des actions orofaciales

Résumé / Abstract

De nombreux modèles récents computationnels et neurobiologiques du contrôle moteur appuient l'hypothèse d'une organisation hiérarchique des représentations d'action au sein du système nerveux central. Ces modèles distinguent différents niveaux de représentations motrices, depuis les représentations musculaires et cinématiques jusqu'aux représentations de plus haut niveau impliquées dans le codage du but des actions. Sur ce point, de nombreuses études sur les primates humains et non humains appuient une localisation du codage du but des actions manuelles au sein du cortex prémoteur et des aires pariétales postérieures. Pour étendre ces résultats aux actions orofaciales, nous avons utilisé dans la présente étude un paradigme de répétition suppression lors de la mesure de l'activité neuronale par imagerie par résonance magnétique fonctionnelle liée à l'exécution répétée de mouvements intransitifs et silencieux des lèvres, de la mâchoire et de la langue. Dans le domaine du contrôle moteur, ce paradigme implique une diminution de l'activité de populations neuronales spécifiques lors de la répétition d'actes moteurs et pourrait refléter des mécanismes d'adaptation sensorimotrice. Dans la présente étude, les mouvements orofaciaux ont activé un ensemble de régions cérébrales communes constituant un réseau neuronal minimal classiquement impliqué dans le contrôle moteur orofacial. De manière cruciale, une diminution des réponses neuronales lors de la répétition des mouvements orofaciaux a été spécifiquement observée dans l'hémisphère gauche, au sein du sulcus intrapariétal et du lobule pariétal inférieur adjacent, du lobule pariétal supérieur et du cortex prémoteur ventral. Ces résultats suggèrent l'existence de mécanismes de codage du but des actions orofaciales intransitives et silencieuses au sein de ce circuit fronto-pariétal.

Recent computational and neurobiological models of motor control argue for a hierarchical organization of action representation in the central nervous system. These models distinguish between different levels of motor representation, from muscle and kinematic representations to higher-level representations involved in action goal coding. From this latter level, numerous studies in both human and nonhuman primates have provided evidence for action goal coding of manual actions in premotor and posterior parietal cortices. To extend these results to orofacial actions, we used a repetition suppression paradigm while measuring neural activity with functional magnetic resonance imaging during repeated intransitive and silent lip, jaw and tongue movements. In the motor domain, this paradigm refers to decreased activity in specific neural populations due to repeated motor acts and has been proposed to reflect sensorimotor adaptation. In the present study, orofacial movements activated a set of largely overlapping, common brain areas forming a core neural network classically involved in orofacial motor control. Crucially, suppressed neural responses during repeated orofacial actions were specifically observed in the left hemisphere, within the intraparietal sulcus and adjacent inferior parietal lobule, the superior parietal lobule and the

ventral premotor cortex. These results provide evidence for action goal coding and motor-to-somatosensory adaptive control of intransitive and silent orofacial actions in this fronto-parietal circuit.

- *Mots clés : Contrôle moteur orofacial, codage du but de l'action, fMRI, répétition suppression, adaptation, sparse sampling.*
- *Keywords: Orofacial motor control, action goal coding, fMRI, repetition suppression, adaptation, sparse sampling.*

2.1.1 Introduction

Codage du but de l'action

Une hypothèse partagée par de nombreux modèles et études portant sur le contrôle moteur est celle d'une organisation hiérarchique de l'action au sein du système nerveux central et impliquant plusieurs niveaux de représentations d'action, depuis les représentations musculaires et cinématiques jusqu'à des niveaux plus abstraits impliqués dans le codage du but et de l'intention de l'action (pour une revue récente, voir Grafton et Hamilton, 2007, qui proposent quatre niveaux de codage de l'action : l'intention, le but de l'action, la stratégie motrice et les paramètres musculaires précis).

En appui de cette décomposition hiérarchique des représentations d'action, des enregistrements unicellulaires dans les aires prémotrices et motrices chez les primates non humains ont fourni des arguments en faveur d'une sélectivité de réponse de populations neuronales spécifiques fonction de différentes variables motrices lors de mouvements manuels (par exemple, force musculaire, vitesse, direction; Kakei, Hoffman et Strick, 1999; Reina et al., 2001; Georgopoulos et al., 1986; Scott et Kalaska, 1997; Scott, Sergio et Kalaska, 1997). A un niveau plus haut d'abstraction dans cette hiérarchie motrice, une sélectivité neuronale pour le codage du but des actions manuelles transitives (soit une action dirigée vers un objet) a été identifiée au niveau des aires postérieures pariétales et prémotrices ventrales (Rizzolatti et al., 1988; Fogassi et al., 2005; Gardner et al., 2007; Rozzi et al., 2008; Umiltà et al., 2001; Bonini et al., 2010, 2011). Notamment, un argument fort en faveur d'un tel codage du but des actions manuelles provient de l'existence de populations neuronales spécifiques au sein des aires pariétales inférieures et prémotrices ventrales montrant des réponses différenciées lors de mouvements de préhension similaires mais dont le but final de l'action diffère (par exemple saisir une cacahuète pour la manger ou pour la déplacer; Fogassi et al., 2005, Bonini et al., 2010).

Chez l'homme, les premiers arguments en faveur d'une organisation du contrôle moteur en différents niveaux de représentations viennent des études neuropsychologiques des patients apraxiques et, notamment, ceux présentant une apraxie idéomotrice qui se caractérise par une incapacité à imiter correctement des gestes manuels et à mimer l'utilisation d'ustensiles malgré une force, une dextérité et une compréhension de la consigne tout à fait correctes. Des travaux pionniers de Liepmann (1920) sur ces patients suggérant un codage des buts d'actions dans le cortex pariétal postérieur, d'autres études ont permis d'associer ce déficit avec des lésions localisées au niveau du sulcus intrapariétal et des aires prémotrices gauches (par exemple, Haaland, Harrington et Knight, 2000). Cette forme d'apraxie peut être contrastée avec une autre forme clinique, l'apraxie idéationnelle (décrite par De Renzi et Lucchelli, 1988), liée à un dysfonctionnement ou une lésion au niveau de la jonction temporo-pariétale gauche, dans laquelle le but d'action est préservé mais le séquençage du mouvement est perturbé.

Plus récemment, les avancées des méthodes de neuroimagerie ont permis d'étudier les corrélats neuronaux liés au codage de l'action chez l'humain sain et de manière non invasive. A titre d'exemple, Johnson-Frey (2004) propose que l'utilisation d'ustensiles (qui nécessite d'une part des

connaissances sémantiques par rapport aux possibles buts d'actions et d'autre part des connaissances procédurales rendant possibles l'exécution de ce but d'action) implique des régions temporales postérieures (pour l'identification de l'objet et les connaissances pragmatiques quant à son utilisation), pariétales postérieures et frontales moyennes/prémotrices gauches (qui permettent de choisir, planifier et exécuter un but d'action correspondant avec l'utilisation de l'objet). En effet, des tâches impliquant la pantomime d'utilisation d'ustensiles, dans lesquelles les buts d'actions tout comme le séquençage et l'exécution doivent être correctement récupérés, ont montré des activations du lobule pariétal inférieur et du gyrus frontal moyen de l'hémisphère gauche (Johnson-Frey et al., 2005). Hamilton et Grafton (2006) ont montré que la partie antérieure du sulcus pariétal inférieur code l'objectif de l'action, ce qui est en accord avec des études plus récentes de Desmurget et al. (2009) qui montrent qu'une stimulation directe du lobule pariétal inférieur induit des envies irrésistibles de mouvoir. En fonction de la force de stimulation, ces envies peuvent même se transformer en véritables illusions de mouvement. Au contraire, la stimulation du cortex prémoteur induit l'effet inverse: les mouvements réels ne sont pas conscientisés, les sujets niant la réalisation effective de leurs propres mouvements. Desmurget et Sirigu (2009) proposent que l'intention motrice soit localisée dans les régions pariétales postérieures inférieures tandis que l'exécution du plan moteur serait liée au cortex prémoteur. Par ailleurs, les auteurs proposent que l'exécution du mouvement peut-être dissociée de la conscience motrice, mais celle-ci apparaîtrait plutôt grâce aux retours des conséquences sensorielles en relation avec les commandes motrices élaborées au sein du cortex prémoteur.

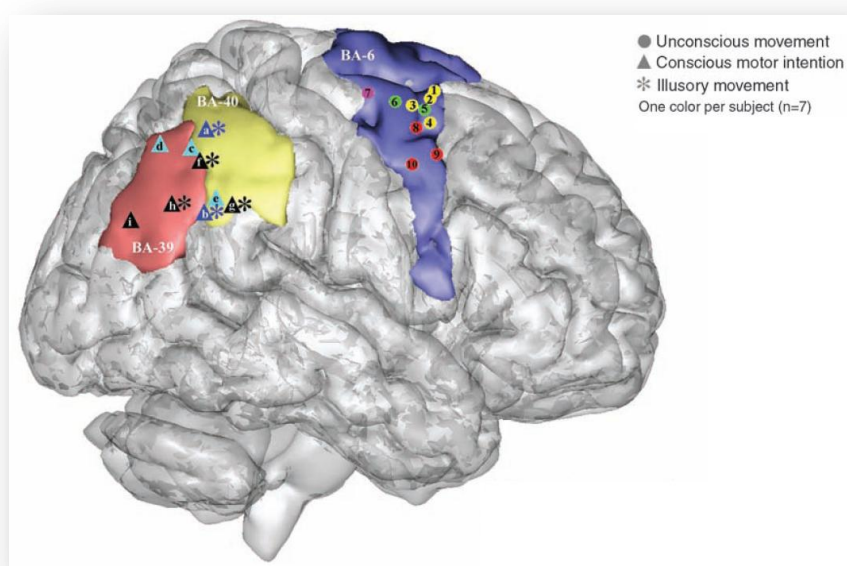


Figure 2.1 : La stimulation des régions pariétales et prémotrices évoque soit l'envie ou l'illusion de mouvements (sans exécution) ou la réalisation effective de mouvements sans conscientisation par le sujet. Les stimulations ont été effectuées chez 9 patients lors d'opérations neurochirurgicales à crâne ouvert (Figure empruntée à Desmurget et al., 2009).

Le mécanisme de Répétition Suppression

Les travaux psychophysiques et comportementaux sur la perception ont démontré que l'observation répétée ou prolongée d'un même type de stimulation induit des modifications perceptives. Cela a été montré par exemple pour la perception visuelle de mouvements (Mather et al., 1998) ou de visages (Webster et al., 2004) ou encore pour la perception auditive de syllabes (Eimas et Corbit, 1973; voir aussi le Chapitre 3.1 de la présente thèse).

Les mécanismes neuronaux sous-jacents à cette adaptation perceptive ont été par la suite étudiés chez le macaque en effectuant des enregistrements unicellulaires. Chez l'homme, la méthode d'imagerie par résonance magnétique fonctionnelle a été pour la première fois utilisée conjointement à un paradigme d'adaptation par Grill-Spector et collègues (1999) afin d'étudier des propriétés invariantes d'objets perçus visuellement. Ce paradigme d'adaptation en IRMf s'est révélé être un outil puissant permettant de dissocier les substrats neuronaux impliqués dans les différents

niveaux de traitements d'un même type de stimulus (Grill-Spector et Malach, 2001; Grill-Spector, Henson et Martin, 2006) et a été récemment également utilisé pour étudier les différents niveaux de représentations motrices des actions manuelles (voir ci-dessous).

Ce paradigme d'adaptation, ou de répétition suppression (RS), est relié à une diminution du signal BOLD lors de l'observation ou de l'exécution répétée de stimuli ou d'actions au sein de régions cérébrales impliquées dans le traitement de l'action observée ou produite⁴⁴. Le mécanisme de RS est étroitement relié à la notion de codage de populations neuronales spécifiques. Un certain nombre de travaux électrophysiologiques ont effectivement démontré un tel codage par des populations neuronales ('population coding') : par exemple, Georgopoulos et al. (1986) ont montré un codage populationnel dans le cortex moteur primaire pour la direction d'un mouvement exécuté tandis que Britten et al. (1993) ont montré un codage de populations neuronales de l'aire visuelle MT dans la discrimination de la direction de mouvement perçus⁴⁵. Lors d'études IRMf, des effets de RS ont été observés pour diverses fonctions et tâches, incluant la vision (Grill-Spector et al., 1999; Kourtzi et Kanwisher, 2000; Grill-Spector et Malach, 2001; Summerfield, 2006), le langage (Joanisse, Zevin et McCandliss, 2007; Menenti et al., 2011, pour plus de détails, voir le Chapitre 2.2), le contrôle moteur (Grafton et Hamilton, 2007; Dinstein et al., 2007; Chong et al., 2008; Hamilton et Grafton, 2009; Lingnau, Gesierich et Caramazza, 2009).

Le paradigme de RS se base sur l'hypothèse que la répétition d'un stimulus ou d'une caractéristique spécifique (trait) de ce stimulus va entraîner une adaptation des populations neuronales codant ce stimulus ou ce trait et une diminution de leurs activités. Ainsi, pour déterminer si une région est sensible à un stimulus ou à un trait donné lors d'une étude IRMf, on présente une suite de stimuli soit différents, soit identiques ou partageant un trait particulier. Par comparaison avec deux stimuli différents, si la répétition du stimulus ou du trait induit une suppression de l'activité d'une région cérébrale (voir Figure 2.2), c'est que les populations neuronales de cette région codent effectivement le stimulus ou le trait commun partagé par les stimuli répétés. Ce paradigme permet ainsi de différencier les régions ou réseaux neuronaux associés à différents types de stimuli ou à différents niveaux de représentation.

Différentes études ont permis de montrer l'existence d'un effet de RS selon différentes échelles temporelles (de quelques millisecondes à quelques minutes voire même des jours; pour une revue, voir Grill-Spector, Henson et Martin, 2006). De plus, cet effet est observable aussi bien au niveau des cellules que du signal BOLD local⁴⁶. Enfin, l'avantage d'utiliser ce paradigme d'adaptation en IRMf consiste également en une augmentation de la spécificité et résolution spatiale des enregistrements.

Il existe plusieurs modèles explicatifs, à ce jour encore débattus, quant aux mécanismes physiologiques reliés à l'effet de RS (voir Figure 2.2) :

- Modèle de fatigue : afin de prévenir une saturation du système et de favoriser la détection de nouveaux stimuli ou traits, la décharge des neurones codant de manière optimale un stimulus ou un trait spécifique est *réduite* lors de leur présentation répétée. De plus, les redondances de codage entre différentes populations sont diminuées de manière à augmenter l'efficacité de l'encodage de l'information.
- Modèle d'aiguïsement ('*sharpening*') : la répétition est considérée comme dépendante d'un

⁴⁴ Il existe également un mécanisme inverse ('repetition enhancement') consistant en une augmentation du signal BOLD (voir Henson et Rugg, 2003).

⁴⁵ Voir aussi les travaux de Hatsopoulos sur l'exploitation du codage populationnel du cortex moteur primaire notamment dans les implémentations cerveau-machine et Ma et Pouget (2009) pour une revue des études théoriques reliées au codage neuronal populationnel.

⁴⁶ De fait, différentes méthodes exploitent ce mécanisme. Par exemple lors d'études EEG/MEG, où différents paradigmes d'adaptation (comme le paradigme oddball classique) permettent de mettre en évidence une modulation des réponses électrophysiologiques lors de la présentation répétée d'un même stimulus ou de stimuli partageant un trait commun. Notamment, la présentation auditive ou visuelle d'un stimulus 'déviant' dans une série de stimuli identiques 'standards' fait apparaître une onde négative autour des 150-250ms (la négativité de discordance, 'Mismatch Negativity' ou MMN; Näätänen, Gaillard et Mäntysalo, 1978; Cammer, 1990).

apprentissage progressif. Au fur et à mesure de la présentation d'un stimulus, les neurones non pertinents au traitement de ce stimulus se désengagent alors que les neurones les plus pertinents renforcent l'encodage. De plus, du fait des interactions compétitives entre neurones, ceux ayant des décharges les plus robustes inhibent l'activité des neurones voisins.

- Modèle de facilitation : ici aussi l'effet de RS est considéré comme dépendant d'un apprentissage progressif. Au fur et à mesure des présentations, une *accélération* des temps de traitements intervient et induit une baisse globale d'activité.

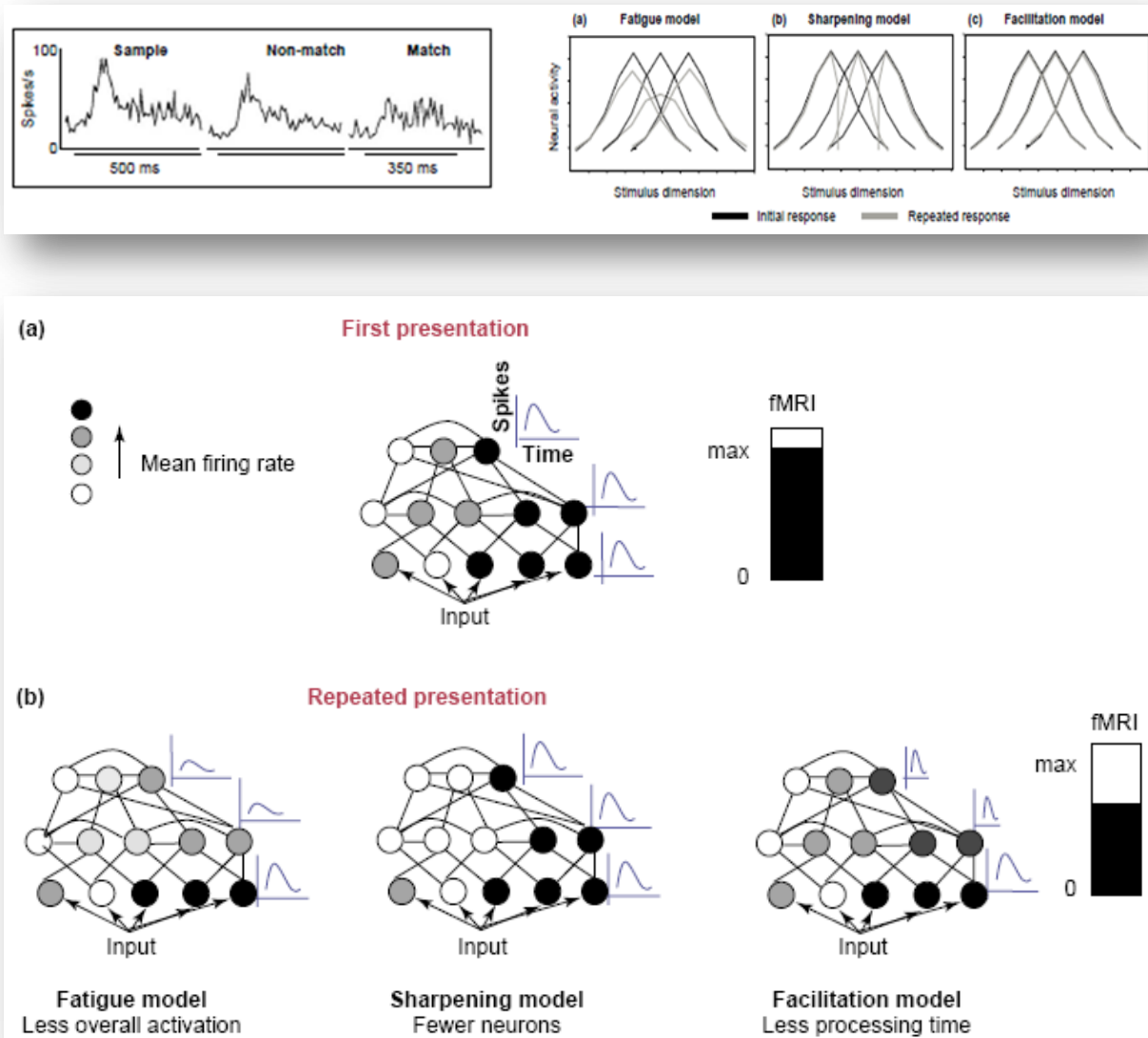


Figure 2.2 : Modèles explicatifs liés à l'effet de RS. En haut à gauche : Décharge d'un neurone du cortex temporal du singe macaque en réponse à la présentation visuelle d'objets (décharges neuronales lors de la première présentation d'un objet, d'un nouveau stimulus et lors de la présentation répétée de l'objet). En haut à droite et en bas : Comparaison des trois modèles de l'effet de RS concernant les modifications de décharge neuronale et d'amplitude du signal BOLD après répétition d'un stimulus. Dans le modèle de fatigue, la réponse de l'ensemble de la population neuronale sensible au stimulus présentée diminue. Dans le modèle d'aiguisement, l'activité des neurones codant des traits non pertinents pour la reconnaissance du stimulus cesse, seuls les neurones pertinents maintiennent leur réponse. Dans le modèle de facilitation, la même population continue à décharger mais le temps de traitement (latence ou durée de la décharge) diminue (Figures empruntées à Grill-Spector, Henson et Martin, 2006).

Répétition suppression pour étudier le but de l'action

A notre connaissance, les corrélats neuronaux du but de l'action chez l'homme ont été étudiés en utilisant le paradigme de RS lors de l'observation ou de l'exécution répétée d'actions manuelles (transitives ou intransitives). En accord avec les études sur les primates nonhumains, ces études ont principalement révélé un effet de RS dans le sulcus intrapariétal et la partie dorsale du lobule pariétal inférieur, ainsi que dans la partie operculaire du gyrus frontal inférieur et le cortex prémoteur ventral adjacent (par exemple, Dinstein et al., 2007; Hamilton et Grafton, 2009; Kilner et al., 2009).

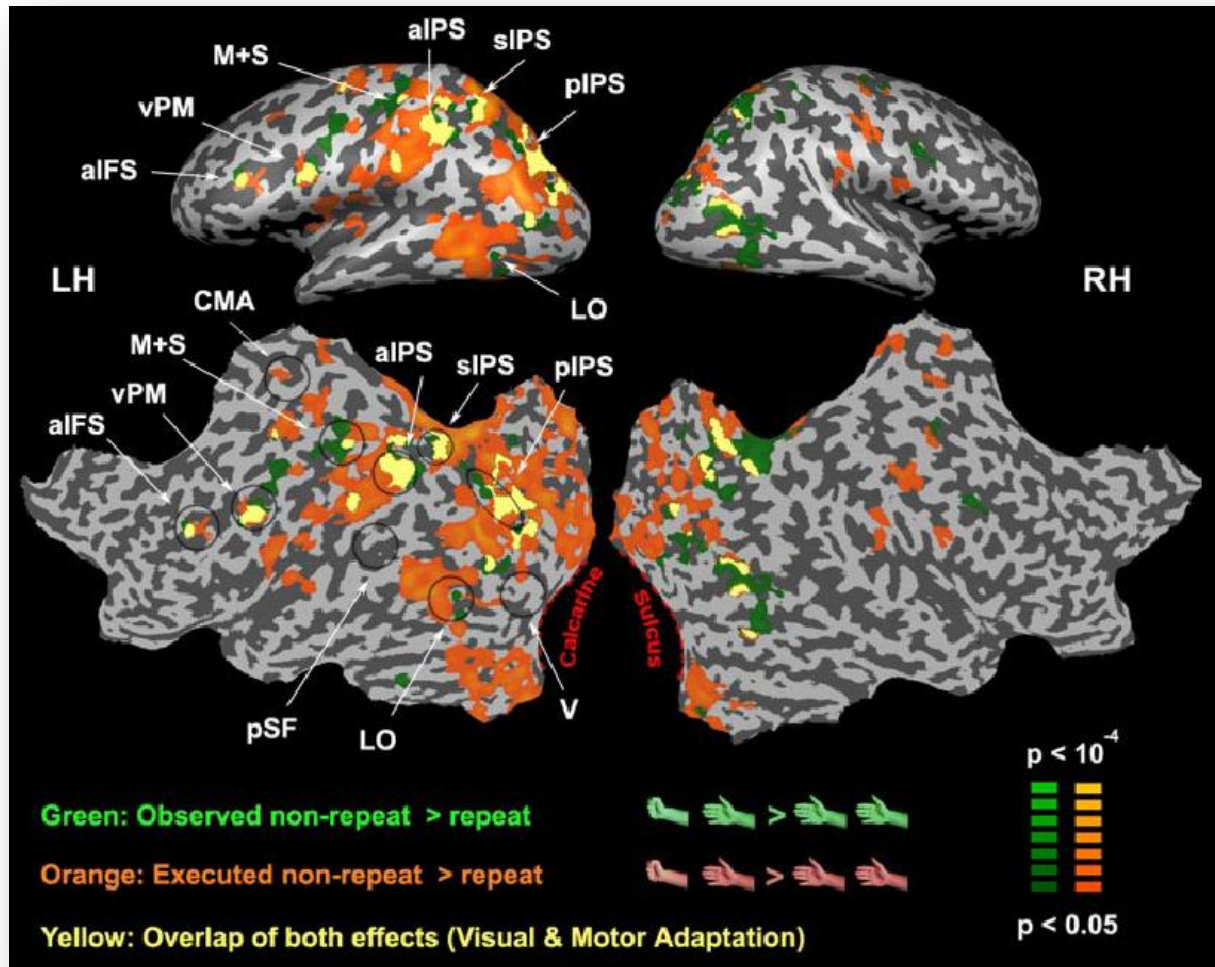


Figure 2.3 : Régions cérébrales sensibles à l'effet de RS lors de l'observation (en vert) ou de l'exécution (en orange) répétée de mouvements manuels intransitifs communicatifs (jeu du "papier-feuille-ciseaux"; Figure empruntée à Dinstein et al., 2007).

Cependant, la plupart de ces études ne montrent pas d'effets d'adaptation cross-modale (voir Hamilton et Grafton, 2006, 2009 et Majdandzic et al., 2009 pour des études impliquant uniquement l'observation ou l'exécution répétée d'une action; ou Dinstein et al., 2007 pour une étude impliquant ces deux tâches mais réalisées séparément) ou alors une adaptation cross-modale asymétrique : soit lors de la perception d'une action suivie de son exécution (Lingnau, Gesierich et Caramazza, 2009), soit lors de l'exécution d'une action suivie par son observation (Chong et al., 2008). Caramazza et collègues (2009) ont exprimé une critique quant à l'interprétation de ce type de résultats en termes de preuve de l'existence d'un système de neurones miroirs chez l'homme impliqué dans le codage du but des actions. En effet, ces auteurs soutiennent qu'il est nécessaire de démontrer une adaptation cross-modale symétrique pour appuyer l'existence d'un système des neurones miroirs. Néanmoins, Kilner et al. (2009) ont depuis démontré un effet d'adaptation au niveau du gyrus frontal inférieur bilatéral aussi bien lors de l'exécution suivie par l'observation d'actions manuelles transitives que lors

de l'observation de ces actions suivies par leur exécution⁴⁷.

Adaptation et processus de contrôle prédictif

Bien que discuté en termes de codage du but de l'action, une interprétation computationnelle convergente de l'effet de RS dans ces aires pariétales et prémotrices est basée sur des processus directs sensorimoteurs adaptatifs. Un modèle direct fait partie du modèle génératif qui consiste en un appariement entre contrôle moteur et conséquences sensorielles dans lequel les schémas de codage prédictif comparent les prédictions top-down avec les informations sensorielles bottom-up, ce qui crée une erreur de prédiction résiduelle, utilisée pour l'estimation et l'apprentissage en ligne (Friston, 2005, 2011). Selon les modèles internes (Wolpert, Ghahramani et Jordan, 1995; Kawato, 1999; Wolpert et Flanagan, 2001), les conséquences sensorielles d'un acte moteur prédites (ou les buts sensoriels) sont générées par le biais des copies d'efférence, en parallèle de la commande motrice. En cas de discordance entre conséquences sensorielles prédites et réelles, un signal d'erreur de prédiction permet ainsi l'apprentissage sensorimoteur et la mise à jour des commandes motrices. En regard des études IRMf d'adaptation sur les actions manuelles, il est possible que les actes moteurs répétés aient causé un apprentissage sensorimoteur graduel et, du fait d'erreurs de prédictions réduites, une mise à jour des représentations motrices de haut niveau impliquées dans le codage du but de l'action dans les aires pariétales et frontales inférieures, ce qui serait reflété par une suppression du signal BOLD. Bien que le codage des buts d'action et les modèles directs aient été largement étudiés dans le contexte des mouvements des membres ou des mains, de tels contrôles sensorimoteurs sont également une idée centrale dans le domaine de la production de la parole (pour une revue récente, voir Perkell, 2012). Par exemple, dans le modèle de production de la parole DIVA (Directions Into Velocities of Articulators; Guenther, 2006; Guenther et Vladusich, 2012), les modulations de réponses des cortex auditif et somatosensoriel sont considérées refléter des mécanismes de contrôle correctif prédictif en ligne, dans lesquels les conséquences sensorielles de l'acte de production de la parole sont évaluées avec l'afférence sensorielle réelle, dans le but de contrôler la production et de distinguer les conséquences sensorielles de nos propres actions des signaux sensoriels causés par des changements de l'environnement (pour des modèles similaires, voir Tian et Poeppel, 2010; Hickok, Houde et Rong, 2011; Price, Crinion et MacSweeney, 2011; voir la partie introductive et la discussion de l'étude 1.2).

Cadre de l'étude

En dehors des études portant sur les processus de contrôle du feedback auditif lors de la production de parole, l'existence de telles boucles régulatrices/prédictives entre systèmes moteur et somatosensoriel lors de l'exécution de mouvements orofaciaux simples, intransitifs et silencieux reste cependant à étudier. Des études précédentes concernant l'exécution de mouvements des lèvres, de la langue ou de la mâchoire/mandibule ont apporté des éléments en faveur de l'existence d'un réseau neural minimal impliqué dans le contrôle moteur orofacial et d'une organisation somatotopique séquentielle dorso-ventrale des articulateurs supralaryngés (voir Chapitre 1.1). En revanche, aucune étude n'a à ce jour étudié les corrélats neuronaux du codage du but de l'action et du contrôle sensorimoteur prédictif lors de mouvements orofaciaux silencieux et intransitifs. Dans la prolongation des résultats mentionnés, la présente étude IRMf avait pour objectif de déterminer une possible suppression de l'activité neurale dans les aires pariétales et prémotrices lors de l'exécution répétée d'actions orofaciales intransitives et silencieuses. Dans ce but, des mouvements de protrusion des lèvres, d'abaissement de la mandibule et de rétraction de la langue ont été exécutés de manière répétitive, lors de séries de six essais consécutifs. De plus, une méthode d'acquisition en

⁴⁷ Il est à noter que les analyses de cette étude étaient restreintes au gyrus frontal inférieur afin d'optimiser la détection des changements du signal BOLD au sein de cette région.

sparse sampling a été utilisée de manière à minimiser de possibles artefacts des images fonctionnelles liés aux mouvements articulatoires.

2.1.2 Méthode

Protocole

Participants

Onze volontaires adultes sains ont participé à l'étude (9 hommes; âge : 20-42 ans, moyenne : 29 ans). Le recrutement des sujets est explicité dans la partie "Considérations Méthodologiques" du premier chapitre.

Procédure expérimentale

L'expérience consistait en la réalisation distincte des tâches suivantes :

- Une protrusion des lèvres (condition "lèvres")
- Une rétraction horizontale de la langue (condition "langue")
- Une ouverture mandibulaire (condition "mâchoire")

Il était demandé aux participants d'initier chacune de ces tâches depuis une position immobile de repos, les muscles linguaux et labiaux relâchés et de revenir à cette position initiale de repos à la fin du mouvement. De manière à évaluer les activations cérébrales spécifiquement associées au contrôle moteur de chacun de ces articulateurs, une condition "contrôle", sans aucune activité motrice, a également été testée.

Les 3 tâches motrices supralaryngées ont été réalisées lors d'un même run fonctionnel (voir partie "Considérations Méthodologiques", Figure 1.1). Durant l'expérience, le sujet devait réaliser l'une des 3 conditions motrices ou la condition de repos toutes les 10 secondes (du fait de l'acquisition des images fonctionnelles de type sparse sampling). Pour caractériser un possible effet de RS, chaque condition était réalisée par trains consécutifs de 6 essais d'une même condition selon un ordre pseudo-aléatoire. Les 3 conditions motrices et la condition de repos consistaient chacune en 3 séries de 18 essais. Pour chaque essai, une consigne visuelle ("lèvres", "langue", "mâchoire" ou "repos") indiquait au sujet la condition à réaliser. Précédée d'une croix de fixation durant 500ms, l'affichage de cette consigne durant 1s servait également de repère temporel aux sujets pour la réalisation de la tâche. Concernant le protocole d'acquisition en sparse sampling, l'intervalle de temps séparant un geste orofacial et le milieu d'acquisition de l'image fonctionnelle correspondante était contrebalancé entre chaque essai et était de 4s, 5s ou 6s. 72 scans fonctionnels ont ainsi été acquis ((3 conditions motrices + 1 condition de repos) x 3 séries x 6 répétitions). 3 scans ont été ajoutés en début de session pour équilibrer le signal IRM et ont ensuite été supprimés des analyses.

Prétraitements des images fonctionnelles

Préalablement aux analyses statistiques, une série de prétraitements des images fonctionnelles a été réalisée avec le logiciel SPM 5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) sous environnement Matlab 7.1 (Mathworks, Natick, MA). Ces prétraitements incluaient des procédures de réalignement, de normalisation dans l'espace commun du Montreal Neurological Institute (repère MNI) et de lissage spatial (voir partie "Considérations Méthodologiques" pour une description détaillée de ces procédures).

Analyses statistiques

Suite aux prétraitements, les données fonctionnelles ont été analysées avec le logiciel SPM 5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) sous environnement Matlab 7.1 (Mathworks, Natick, MA). Les régions cérébrales activées ont été labellisées en utilisant la toolbox SPM Anatomy (Eickhoff et al., 2005) et, si nécessaire, en utilisant le logiciel Talairach Daemon (Lancaster et al., 2000). Les analyses par région d'intérêt (ROI) ont été effectuées via la toolbox SPM Anatomy. Pour la visualisation des résultats, les cartes d'activation ont été superposées sur un template anatomique standard en utilisant le logiciel MRICRON (<http://www.sph.sc.edu/comd/rorden/mricron/>).

Analyses individuelles

Au niveau individuel et suite aux prétraitements des images fonctionnelles (voir partie "Considérations Méthodologiques"), les activations neuronales liées aux gestes orofaciaux supralaryngés ont dans un premier temps été estimées par l'utilisation d'un Modèle Linéaire Général (GLM; Friston et al., 1995) et ce, indépendamment des différentes répétitions de chacune des actions orofaciales. Pour chaque participant, un second GLM a été estimé de manière à évaluer les activités reliées à chacune des 6 répétitions (1 à 6) ce, indépendamment des tâches motrices. Ces deux analyses n'ont pu être regroupées au sein d'une même analyse du fait d'un trop faible nombre d'essais par conditions (3 conditions x 6 répétitions x 3 scans).

Les modèles incluaient des régresseurs d'intérêt reliés soit aux 3 tâches motrices (chacune représentée par 18 images fonctionnelles) indépendamment des répétitions, soit aux 6 répétitions indépendamment des tâches motrices (chacune représentée par 9 images fonctionnelles) et des régresseurs de non-intérêt liés aux paramètres de réalignement. Les conditions de repos formaient une ligne de base.

Pour les deux modèles, la réponse de type hémodynamique associée à chaque événement a été modélisée par une fonction à réponse impulsionnelle finie et à impulsion unique (FIR) pour chaque scan fonctionnel. Avant l'estimation du modèle, un filtrage des basses fréquences *a priori* non reliées aux conditions expérimentales (variations lentes d'origine physiologique) a été appliqué (passe-haut de fréquence de coupure de 1/128 Hz). Les estimations des réponses FIR modélisées ont été calculées de manière à correspondre avec le décours temporel du signal BOLD observé pour chaque voxel et chaque condition. Pour chaque participant, des cartes d'analyse statistique individuelles ont ainsi été calculées pour chacune des 3 tâches motrices dans le premier modèle et pour chacune des 6 répétitions dans le second modèle, par contraste avec la condition contrôle. Ces cartes statistiques individuelles ont été ensuite utilisées pour les analyses de groupe.

Analyses de groupe

Afin de permettre une généralisation des résultats observés pour la population étudiée (Friston, Holmes et Worsley, 1999), des analyses de groupe de second niveau à effet aléatoire ont été réalisées au moyen d'analyses de variance (ANOVA) à mesures répétées.

Mouvements orofaciaux

De manière à déterminer les activités cérébrales liées aux mouvements orofaciaux, une ANOVA à mesures répétées a été réalisée avec comme facteur intra-sujet la tâche motrice (incluant 3 niveaux liés aux mouvements des lèvres, de la langue et de la mandibule). La variable "sujets" a été prise en compte comme facteur aléatoire.

Trois contrastes t unidirectionnels ont été calculés afin de déterminer les régions cérébrales activées de manière spécifique pour chacune des tâches orofaciales par rapport à la condition de repos. Dans le but d'identifier les activations communes/partagées par toutes les conditions motrices, une analyse de conjonction (Friston et al., 1999; Nichols et al., 2005) a également été effectuée. Finalement, un contraste F bidirectionnel correspondant à l'effet principal de la condition motrice a été calculé dans le but de déterminer les différences d'activations cérébrales entre les trois tâches motrices.

Effet de RS

Une seconde analyse de groupe a été réalisée dans le but de comparer les activations cérébrales fonctions des 6 répétitions, indépendamment des tâches motrices supralaryngées. Pour ce faire, une seconde ANOVA à mesures répétées a été effectuée avec comme facteur intra-sujet l'ordre de répétition (incluant 6 niveaux). La variable "sujets" a été prise en compte comme facteur aléatoire.

Six contrastes t ont été calculés de manière à déterminer les régions cérébrales activées de manière spécifique pour chacune des six répétitions par rapport à la ligne de base. Dans le but d'identifier un possible effet de RS, un contraste t représentant une décroissance progressive linéaire des activations en fonction des 6 répétitions a été calculé.

Somatotopie sensorimotrice

Des analyses par régions d'intérêt (ROI) ont été effectuées afin de tester une possible organisation somatotopique des activations supralaryngées et laryngées au sein des structures primaires motrice et somatosensorielle. Comme pour l'étude 1.1, trois régions d'intérêt spécifiques correspondant aux cortex primaire moteur et somatosensoriel ont été créées en utilisant la toolbox SPM Anatomy (Eickhoff et al., 2005; voir Figure 2.6). La première ROI correspondait au cortex moteur primaire et était composée des cartes cytoarchitectoniques des aires de Brodmann BA4a et BA4p (parties antérieures et postérieures du cortex moteur primaire; voir Geyer et al., 1996). Etant données les positions anatomiques des régions du cortex somatosensoriel primaire BA3 et BA1 (partie intérieure du sulcus central et partie postérieure du gyrus postcentral respectivement; voir Geyer et al., 1999), deux ROIs indépendantes ont été créées pour analyser les activités du cortex somatosensoriel primaire correspondant aux cartes cytoarchitectoniques des aires 3a et 3b combinées pour correspondre à l'aire de Brodmann BA3 et à la carte cytoarchitectonique de l'aire BA1.

Les analyses par ROI ont été réalisées individuellement pour chaque participant et chaque tâche, en appliquant une correction de type "small volume correction" pour chaque ROI selon un seuil statistique défini à $p < .001$ non corrigé. Les coordonnées MNI du pic d'activation maximale et du centre de gravité (COG) au sein des cortex primaires moteur et somatosensoriel (BA4, BA3 et BA1) ont été déterminées dans les axes médio-latéral (X), rostro-caudal (Y) et dorso-ventral (Z) et ce pour chaque participant, chaque hémisphère et pour toutes les conditions motrices. Pour chaque ROI, des ANOVA à mesures répétées à deux facteurs ont été calculées. Les facteurs intra-sujet impliquaient l'hémisphère (2 niveaux : gauche, droite) et l'articulateur (3 niveaux : lèvres, langue, mandibule). Les analyses ont été effectuées pour les dimensions x, y et z indépendamment (en valeurs absolues des coordonnées MNI pour l'axe x). Pour ces analyses le niveau de significativité a été défini à $p < .05$ et une correction Greenhouse-Geisser a été appliquée en cas de non sphéricité des données. Lorsque nécessaire, des analyses post-hoc ont été effectuées en utilisant des tests de type LSD Fischer.

2.1.3 Résultats

Les résultats des analyses de groupe liées aux réseaux neuronaux des mouvements orofaciaux et des répétitions sont reportés selon un seuil statistique défini à $p < .05$ après correction de type "Family-Wise Error" (FWE; Nichols et Hayasaka, 2003) et une taille minimale des clusters de 30 voxels,

excepté pour l'effet principal de la tâche motrice et l'effet de RS (seuil défini à $p < .001$ non corrigé et une taille minimale des clusters de 30 voxels). Pour tous les contrastes, les pics d'activation maximum ont été déterminés dans chaque cluster, leur localisation a ensuite été labellisée avec la boîte à outils Anatomy de SPM (Eickhoff et al., 2005). Si une région n'avait pas pu être assignée avec Anatomy, elle a été déterminée avec le logiciel Talairach Daemon (Lancaster et al., 2000) grâce aux coordonnées du pic d'activation converties de l'espace MNI à l'espace stéréotaxique standard de Talairach et Tournoux (1988).

Réseaux neuronaux des mouvements orofaciaux

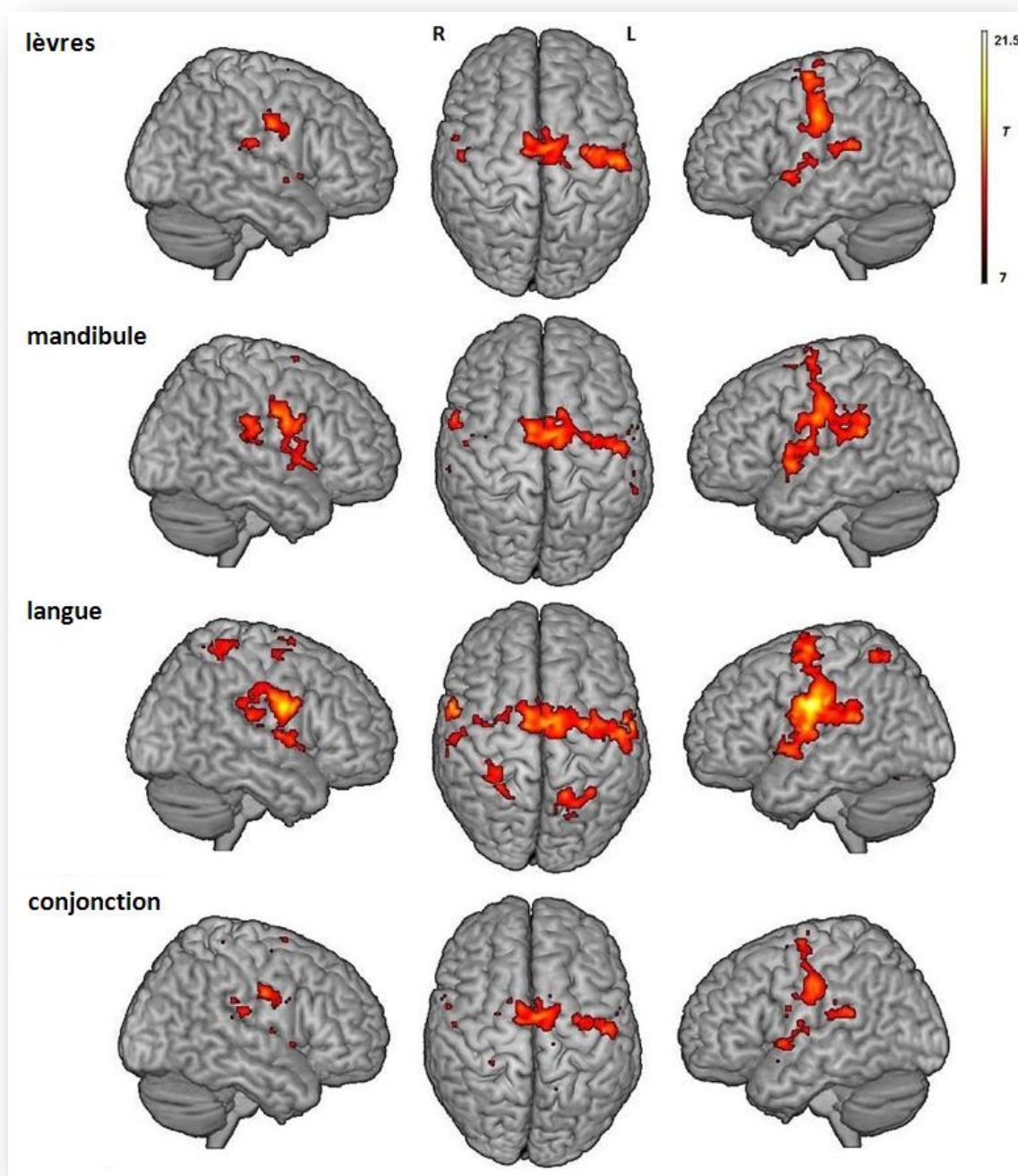


Figure 2.4 : Réseaux cérébraux des mouvements labiaux, mandibulaires et lingaux et réseau minimal du contrôle moteur orofacial tel que révélé par l'analyse de conjonction. Résultats reportés à un seuil corrigé (FWE, $p < .05$) et une taille minimale des clusters de 30 voxels.

Réseau articulatoire commun

Les activations cérébrales observées pour les trois tâches motrices, pour les analyses de conjonction et de différences entre tâches sont présentées dans la Table 2.1 et la Figure 2.4. Les résultats de l'analyse de groupe montrent des régions largement communes aux trois tâches motrices. L'analyse de conjonction révèle en effet un réseau neuroanatomique fonctionnel commun incluant un ensemble de régions bilatérales typiquement impliquées dans le contrôle moteur orofacial. Ce réseau orofacial 'minimal' (voir Chapitre 1.1) implique des activations bilatérales de l'aire motrice supplémentaire et des cortex prémoteur et sensorimoteur (incluant l'opercule pariétal). Des activations sont également observées dans l'hémisphère gauche, au niveau de l'insula et du claustrum, du striatum dorsal des ganglions de la base (putamen), du gyrus temporal transverse et du cortex pariétal inférieur, et dans la partie supérieure du cervelet (région déclive du néocervelet).

Régions	H	BA	Coordonnées MNI			T
			x	y	z	
Cluster 1 (519 voxels)						
Aire motrice supplémentaire	L	6	-6	-6	56	13.42
Aire motrice supplémentaire	R	6	8	-8	58	12.91
Lobule paracentral	L	6	-14	-12	74	7.45
Cluster 2 (457 voxels)						
Cortex sensorimoteur primaire	L	4	-44	-16	38	12.04
Gyrus frontal moyen	L	6	-30	-8	68	10.34
Cortex pariétal inférieur	L	40	-54	-22	34	9.68
Cortex prémoteur ventral	L	6	-56	-4	32	8.89
Cortex prémoteur dorsal	L	6	-46	-10	52	8.83
Cluster 3 (267 voxels)						
Cortex sensorimoteur primaire	R	4	50	-8	34	12.61
Cortex prémoteur ventral	R	6	38	-12	38	11.62
Cortex préfrontal	R	9	54	0	26	7.92
Cluster 4 (198 voxels)						
Insula	L	13	-44	0	2	10.79
Cortex prémoteur	L	6	-48	-4	8	9.15
Opércule pariétal	L	43	-60	-6	14	8.97
Cluster 5 (172 voxels)						
Insula	L	13	-52	-40	20	11.01
Cortex pariétal inférieur	L	40	-54	-36	22	10.59
Gyrus temporal transverse	L	41	-36	-26	8	10.05
Opércule pariétal	L	43	-46	-30	18	8.48
Cluster 6 (115 voxels)						
Clastrum	L		-32	-4	-8	11.19
Putamen	L		-24	-6	-4	9.96
Cluster 7 (61 voxels)						
Insula	R	13	46	6	0	10.99
Cluster 8 (38 voxels)						
Opércule pariétal	R	43	60	-22	22	9.45
Cluster 9 (33 voxels)						
Cervelet (declive)	L		-18	-60	-22	9.33

Table 2.1 : Coordonnées dans l'espace MNI des pics d'activation maximale commun aux 3 actions orofaciales (analyse de conjonction). Résultats reportés à un seuil corrigé (FWE, $p < .05$) et une taille minimale des clusters de 30 voxels.

Activations motrices fonction de l'articulateur

En dehors du réseau minimal des tous les articulateurs et des différences observées dans l'analyse de l'effet principal, des activations supplémentaires ont été reportées pour les mouvements linguaux, les activations labiales et mandibulaires (voir Figure 2.4). Pour les mouvements linguaux, des

activations additionnelles ont été trouvées bilatéralement au sein du lobule pariétal supérieur incluant l'aire somatosensorielle associative et le précuneus gauche (BA5 et BA7) et le culmen gauche et l'amygdale droite. Comme lors de notre précédente étude (voir Chapitre 1.1), les étendues des clusters activés dans la tâche linguale ont été significativement plus étendues au sein du cortex sensorimoteur par rapport aux autres mouvements orofaciaux. Ces activations n'apparaissant ni dans l'analyse de conjonction ni dans l'analyse de l'effet principal, ces résultats sont cependant à considérer avec précaution.

Effet principal de la tâche motrice

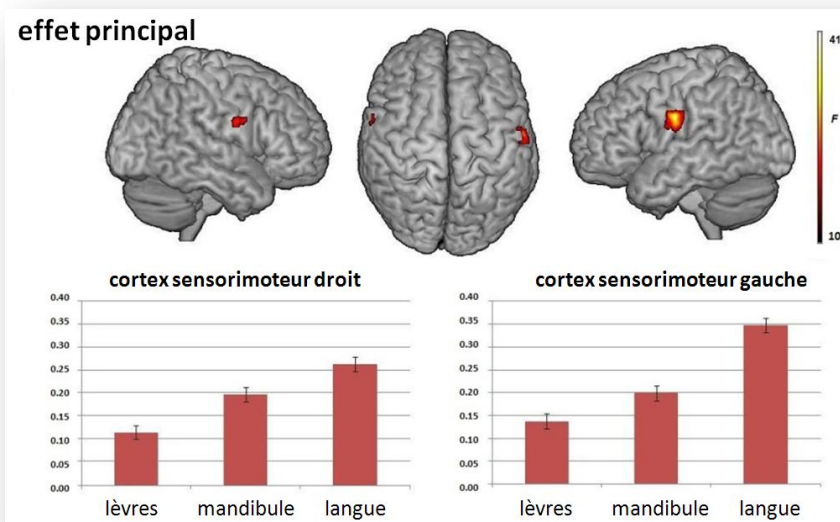


Figure 2.5 : Effet principal de la tâche motrice. Haut: vue surfacique des régions cérébrales activées de manière différenciée pour la condition linguale par rapport aux autres tâches. Bas: contrastes β , reflétant le pourcentage de changement du signal BOLD par rapport à la condition de repos pour les régions activées et pour les 3 tâches motrices ($p < .001$, non corrigé, et taille minimale des clusters de 30 voxels).

Des différences d'activations entre les trois tâches motrices sont observées au niveau de régions du cortex prémoteur ventral (BA6, x/y/z : -52/-4/20 et 54/-2/24) et sensorimoteur primaire gauche (BA3, x/y/z : -56/-10/28). Ces différences proviennent toutes d'une activation plus importante pour les mouvements de la langue et moindre pour les mouvements de lèvres (voir Figure 2.5).

Somatotopie sensorimotrice

Les analyses par ROI ont été réalisées individuellement à l'aide de la toolbox SPM Anatomy pour chaque participant et chaque tâche, en appliquant une correction de type "small volume correction" pour chaque ROI selon un seuil statistique défini à $p < .001$ non corrigé. Les coordonnées MNI du pic d'activation maximale et du centre de gravité (COG) au sein des cortex primaire moteur et somatosensoriel (BA4, BA3 et BA1) ont été déterminées dans les axes médio-latéral (X), rostro-caudal (Y) et dorso-ventral (Z) et ce pour chaque participant individuellement, dans les deux hémisphères et pour toutes les conditions motrices (voir Figure 2.6). Les données d'un sujet ont été écartées de cette analyse, car le seuil choisi n'a pas permis de faire apparaître des activations sensorimotrices pour toutes les tâches et les deux hémisphères.

Pour chaque ROI, des ANOVA à mesures répétées à deux facteurs ont été calculées. Les facteurs intra-sujet impliquaient l'hémisphère (2 niveaux : gauche, droite) et l'articulateur (4 niveaux : lèvres, langue, mandibule, larynx). Les analyses ont été effectuées pour les dimensions x, y et z indépendamment (en valeurs absolues des coordonnées MNI pour l'axe x). Pour ces analyses le niveau de significativité a été défini à $p < .05$ et une correction Greenhouse-Geisser a été appliquée en cas de non sphéricité des données. Lorsque nécessaire, des analyses post-hoc ont été effectuées en utilisant des tests de type LSD Fischer.

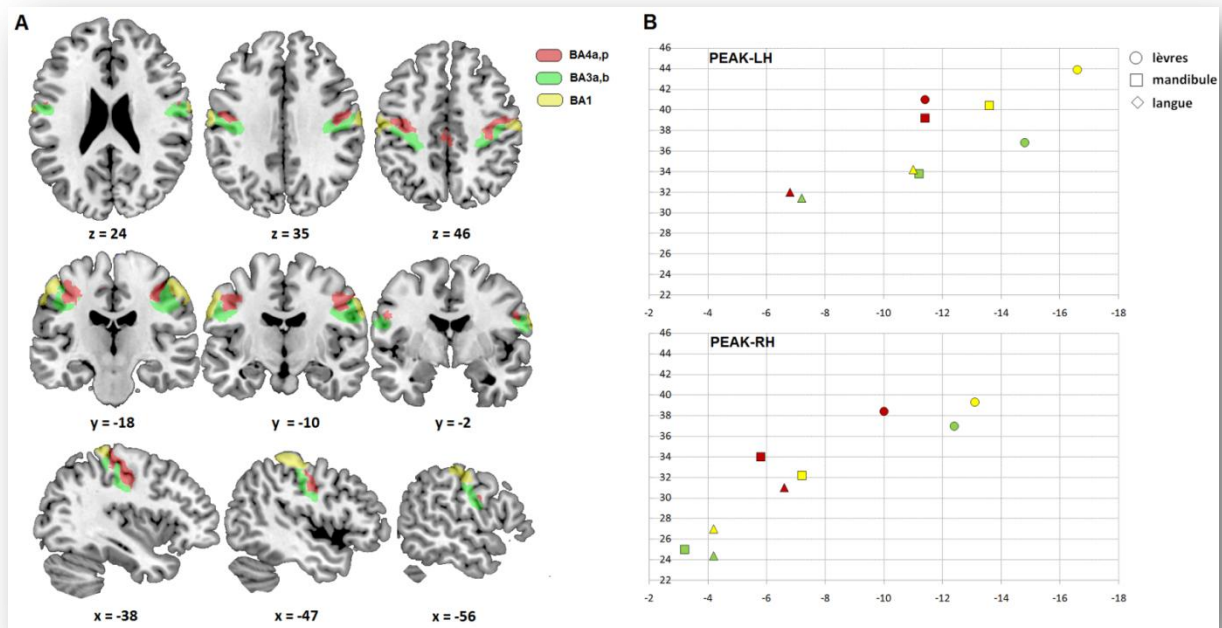


Figure 2.6 : Somatotopie sensorimotrice (analyse par régions d'intérêts). A) ROI correspondant aux aires BA 4a/4p (rouge), BA 3a/3b (vert) et BA 1 définies via la Toolbox SPM Anatomy. B) Projections des coordonnées Y et Z des pics d'activation maximale des trois articulateurs. Une organisation somatotopique est observée, impliquant une position plus postérieure des lèvres par rapport à la langue dans la dimension rostro-caudale (Y) et une position plus dorsale des lèvres par rapport à la langue dans l'axe dorso-ventral (Z).

Pour toutes les ROIs sensorimotrices, une position plus dorsale des lèvres par rapport à la mandibule et la langue est observée. Un effet significatif de l'articulateur a été également trouvé pour les ROI BA3 et BA1 dans l'axe médio-latéral, avec une position plus latérale de la langue que des lèvres. Enfin, pour BA3 seulement, une interaction entrehémisphère et articulateur est observée (avec une position plus ventrale de la mandibule dans l'hémisphère droit).

Dimension médio-latérale (x)

- **BA4:** Aucun effet significatif n'est observé.
- **BA3:** Un effet significatif du facteur articulateur est observé ($F_{(2,18)} = 5.73$; $p < .01$) avec une position plus latérale pour la langue (53cm) par rapport à la mandibule (50cm) et aux lèvres (47cm).
- **BA1:** Des effets significatifs des facteurs Hémisphère et Articulateur sont observés ($F_{(1,10)} = 21.9$; $p < .001$; $F_{(3,30)} = 3.8$; $p < .02$), avec une position plus médiale dans l'hémisphère gauche (53 vs 57mm) et une position plus latérale pour la langue par rapport aux lèvres et à la mandibule (58, 53, 54mm, respectivement).

Dimension rostro-caudale (y)

- **BA4:** Aucun effet significatif n'est observé.
- **BA3:** Des effets significatifs des facteurs Hémisphère et Articulateur sont observés ($F_{(1,10)} = 5.54$; $p < .034$; $F_{(2,18)} = 10.08$; $p < .001$) avec une position globale plus postérieure dans l'hémisphère gauche (-13 vs -8mm) et une position plus postérieure pour les lèvres (14mm) que pour la mandibule (9mm) et la langue (7mm).
- **BA1:** Aucun effet significatif n'est observé.

Dimension dorso-ventrale (z)

- **BA4**: Un effet significatif du facteur Articulateur est observé ($F_{(2,18)} = 6.69$; $p < .007$), avec une position plus dorsale pour les lèvres (40mm) et la mandibule (37mm) par rapport à la langue (32mm).
- **BA3** : Un effet significatif du facteur Articulateur est observé ($F_{(2,18)} = 10.51$; $p < .001$), avec une position plus dorsale pour les lèvres (37mm) par rapport à la mandibule (29mm) et la langue (28mm).
- **BA1**: Un effet significatif du facteur Articulateur est observé ($F_{(2,18)} = 8.13$; $p < .003$), avec une position plus dorsale pour les lèvres (42mm) par rapport à la langue (31mm).

Effet de Répétition Suppression

Les projections des activations cérébrales observées pour les six répétitions ainsi que l'effet de RS sont présentés dans la Table 2.2 et la Figure 2.7.

Régions	p cluster	H	BA	Coordonnées MNI			T	Contrastes estimés					
				x	y	z		1	2	3	4	5	6
Cluster 1 (271 voxels)	0.000												
Cortex pariétal inférieur		L	40	-58	-38	42	5.15	0.14	0.13	0.09	0.06	0.01	0.06
Sulcus intrapariétal		L		-40	-42	36	4.44	0.11	0.08	0.05	0.03	0.02	0.03
Cluster 2 (173 voxels)	0.004												
Cortex prémoteur		L	6	-34	2	26	4.50	0.08	0.01	0.00	-0.02	-0.01	-0.04
Cluster 3 (108 voxels)	0.042												
Cortex pariétal supérieur		L	7	-28	-64	54	4.08	0.11	0.05	0.01	0.07	-0.02	0.00
Cortex pariétal inférieur		L	40	-32	-50	44	3.74	0.15	0.07	0.04	0.06	0.03	0.06
Cluster 4 (73 voxels)	0.160												
Gyrus cingulaire			32	0	16	42	4.05	0.17	0.11	0.04	0.08	0.01	0.01
Cluster 5 (71 voxels)	0.173												
Cortex prémoteur dorsal		L	6	-32	-10	50	4.48	0.10	0.05	0.01	0.08	0.04	-0.01
Cluster 6 (57 voxels)	0.301												
Sulcus intrapariétal		R		36	-52	46	4.27	0.08	0.05	0.03	0.03	-0.01	-0.02
Cortex pariétal supérieur		R	7	34	-50	54	3.56	0.10	0.04	0.01	0.08	0.01	-0.01
Cluster 7 (53 voxels)	0.351												
Cortex extra-strié		L	19	-46	-56	-8	3.90	0.09	0.01	0.00	0.00	-0.01	-0.01
Gyrus fusiforme		L	37	-54	-56	0	3.74	0.12	0.08	0.06	0.03	0.01	0.05
Cluster 8 (37 voxels)	0.622												
Gyrus frontal inférieur		R	47	42	30	2	3.66	0.09	0.08	0.05	0.04	0.02	0.02
Cluster 9 (35 voxels)	0.662												
Gyrus frontal inférieur		L	47	-36	30	-4	3.88	0.12	0.02	0.04	0.02	0.00	0.00

Table 2.2 : Coordonnées dans l'espace MNI des régions montrant un effet de RS (décroissance du signal BOLD) pour les 6 productions répétées des mouvements orofaciaux. Résultats reportés à un seuil non corrigé $p < .001$, excepté pour le lobules pariétaux inférieur et supérieur et le cortex prémoteur ventral de l'hémisphère gauche avec correction au niveau du cluster, et une taille minimale des clusters de 30 voxels.

Les principales régions montrant un effet de RS sont observées dans l'hémisphère gauche au niveau du lobule pariétal inférieur et du sulcus intrapariétal ainsi que dans le lobule pariétal supérieur (précuneus) et la partie la plus dorsale du cortex prémoteur ventral. Ces résultats sont cohérents avec les précédentes études montrant l'encodage dans ces régions des buts d'actions manuelles (actions avec retours sensoriels somatosensoriels et visuels). D'autres régions supplémentaires montrent également un effet de RS (le sulcus intrapariétal et le précuneus droit, le gyrus cingulaire antérieure, le gyrus frontal moyen gauche et des régions visuelles gauches, comme le gyrus fusiforme et le cortex extrastrié), sans avoir cependant survécu à un seuil corrigé au niveau du cluster.

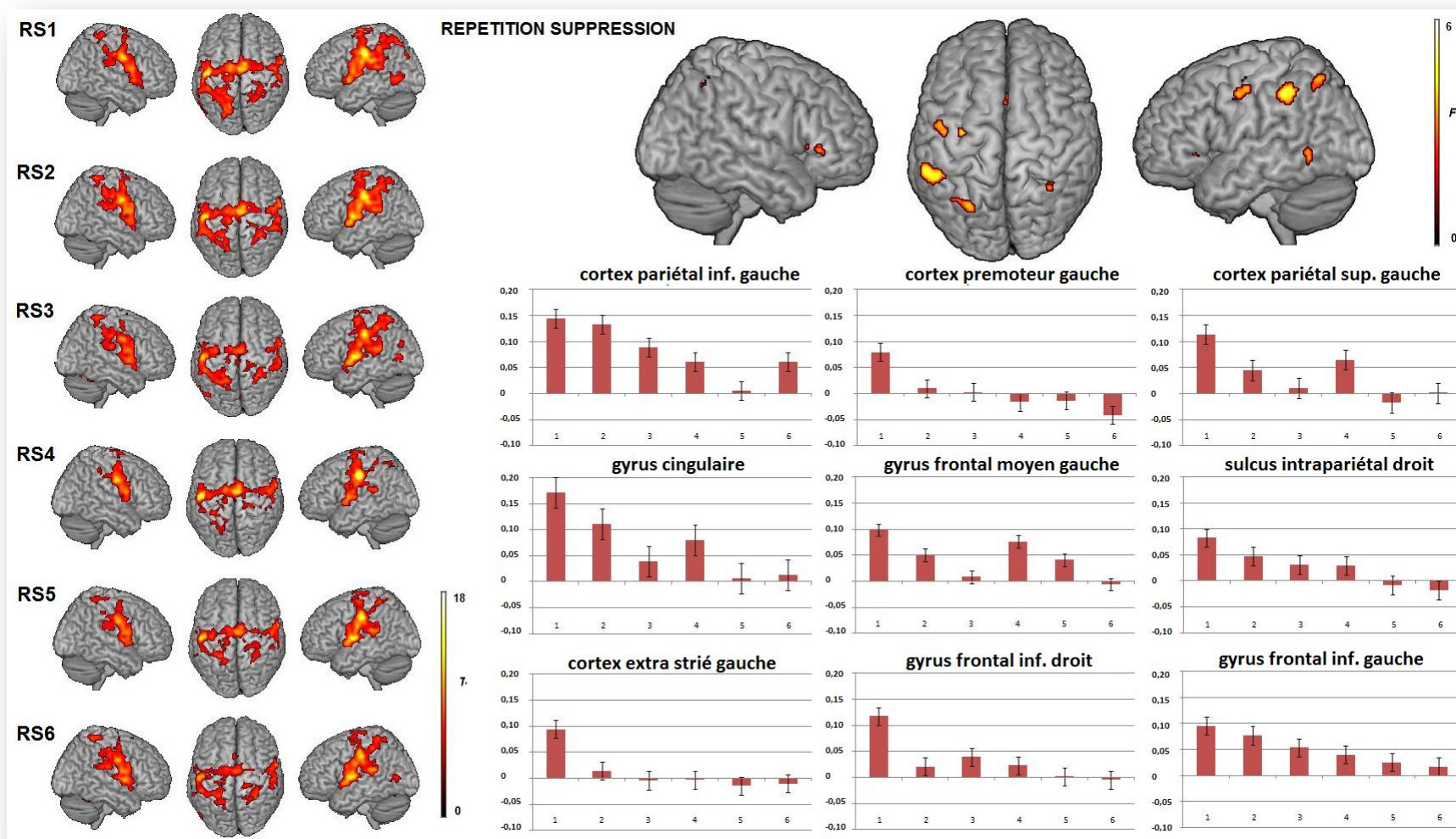


Figure 2.7. Gauche : Vues surfaciques sur les régions cérébrales activées lors de chaque occurrence répétée d'un même geste orofacial (tous gestes labial, mandibulaire et lingual confondus). Droite : Régions cérébrales sensibles à un effet de répétition suppression avec les estimations des contrastes β , reflétant le pourcentage du changement du signal BOLD à partir de la ligne de base pour les six occurrences répétées ($p < .001$, non corrigé, excepté pour les lobules pariétaux inférieur et supérieur et le cortex prémoteur ventral de l'hémisphère gauche avec correction au niveau du cluster, et taille minimale des clusters de 30 voxels)

2.1.4 Discussion

En utilisant un protocole d'adaptation, le but de cette étude IRMf en sparse sampling était de déterminer un possible effet de répétition suppression lors de mouvements silencieux et intransitifs des lèvres, de la langue et de la mandibule. De manière générale, ces mouvements orofaciaux ont activé un ensemble de régions largement communes classiquement impliquées dans le contrôle moteur orofacial. De manière cruciale, une diminution des réponses neurales lors de l'exécution répétée des actions orofaciales a été spécifiquement observée dans l'hémisphère gauche, au sein du sulcus intrapariétal et du lobule pariétal inférieur adjacent, du lobule pariétal supérieur et du cortex prémoteur ventral. Ces résultats arguent en faveur de processus de codage du but des actions orofaciales silencieuses et intransitives et de mécanismes de contrôle prédictif motorio-somatosensoriel dans ce circuit fronto-pariétal.

Actions orofaciales

Comme mis en évidence par l'analyse de conjonction (voir Figure 2.4 et Table 2.1) et non spécifiquement relié aux effets de RS, les mouvements labiaux, linguaux et mandibulaires ont activé un ensemble de régions cérébrales communes, incluant le cortex sensorimoteur et prémoteur, l'aire motrice supplémentaire, l'opercule pariétal, l'insula, les ganglions de la base à gauche, le claustrum et le cervelet gauche. Ces résultats apparaissent parfaitement concordants avec de précédentes études IRMf portant sur le contrôle moteur orofacial et une fonction globale des régions activées dans des processus de préparation, d'exécution et de régulation motrice (Lotze et al., 2000; Hesselmann et al., 2004; Riecker et al., 2005; Dhanjal et al., 2008; Grabski et al., 2012; voir Chapitre 1.1). Notamment, ces résultats confirment ceux observés lors de notre précédente étude portant sur le réseau minimal de contrôle moteur des mouvements supralaryngés et laryngés, utilisant un paradigme expérimental en tous points identique mais sans l'utilisation d'un protocole d'adaptation (voir Chapitre 1.1). En dehors de ces activations communes aux trois gestes supralaryngés, l'analyse de l'effet principal révèle des différences d'activité significatives entre les trois mouvements orofaciaux au niveau du cortex prémoteur ventral et sensorimoteur primaire gauche et du cortex prémoteur ventral droit. Toutes ces différences étant dues à une activité accrue pour les mouvements de rétraction de la langue, ce résultat pourrait refléter (1) des demandes de contrôle moteur plus complexes pour la langue par rapport aux autres articulateurs dans ces régions et/ou (2) la localisation de la langue par rapport aux autres articulateurs au sein d'une somatotopie sensorimotrice. Quelques études précédentes ont en effet montré des clusters d'activation dans le gyrus précentral plus étendus pour les mouvements linguaux que pour des mouvements labiaux et/ou mandibulaires (Hesselmann et al., 2004; Pulvermüller et al., 2006; Dhanjal et al., 2008; Grabski et al., 2012) ainsi qu'une localisation ventrale similaire au sein du cortex sensorimoteur pour les mouvements linguaux (pour une revue, voir Chapitre 1.1).

Répétition Suppression

De manière cruciale, un effet de RS au travers des six occurrences successives des actions orofaciales a été observé dans le lobule pariétal inférieur dorsal gauche et le sulcus intrapariétal adjacent, le lobule pariétal supérieur gauche et le cortex prémoteur ventral gauche⁴⁸ (voir Figure 2.7 et Table

⁴⁸ D'autres régions supplémentaires montrent une sensibilité à la répétition des mouvements orofaciaux (le sulcus intrapariétal et le précuneus droit, le gyrus cingulaire antérieure, le gyrus frontal moyen gauche et des régions visuelles gauches, comme le gyrus fusiforme et le cortex extra strié), sans avoir cependant survécu à un seuil corrigé au niveau du cluster. Le cortex cingulaire antérieur dorsal et le cortex prémoteur dorsal sont connues pour jouer un rôle dans les fonctions motrices de haut niveau telles que l'attention motrice et la sélection de mouvement (Devinsky et al., 1995;

2.2). Pris ensemble, ces résultats apparaissent en accord avec des études portant sur le codage du but des actions manuelles réalisées chez les primates humains et non humains. Comme mentionné précédemment, des enregistrements neurophysiologiques unicellulaires chez le primate non humain ont montré une sélectivité neuronale pour le codage du but des actions manuelles dans les aires pariétales postérieures et prémotrices ventrales (Rizzolatti et al., 1988; Fogassi et al., 2005; Gardner et al., 2007; Rozzi et al., 2008; Umiltà et al., 2008; Bonini et al., 2010, 2011). Chez les humains, les premiers arguments en faveur du codage des buts d'actions dans le cortex pariétal postérieur proviennent des travaux pionniers de Liepmann (1920/1988) et des études ultérieures sur les patients souffrant d'apraxie idéomotrice. L'apraxie idéomotrice se caractérise chez ces patients par une incapacité à imiter correctement des gestes manuels et à mimer l'utilisation d'ustensiles, un déficit essentiellement associé avec des lésions localisées au niveau du sulcus intrapariétal et des aires prémotrices gauches (par exemple, Haaland, Harrington et Knight, 2000), et ce malgré une force, dextérité et compréhension tout à fait correctes.

Chez l'homme, des études IRMf utilisant un protocole d'adaptation ont également révélé que les actions manuelles répétées transitives et intransitives communicatives (Dinstein et al., 2007; Hamilton et Grafton, 2009; Kilner et al., 2009) avec des buts similaires induisent un effet de RS au niveau du sulcus intrapariétal antérieur et au niveau du gyrus frontal inférieur et du cortex prémoteur. De manière intéressante, ces effets de RS observés dans les aires postérieures pariétales et prémotrices ne sont pas restreints aux actes moteurs exécutés mais apparaissent tout aussi bien lors de l'observation répétée d'actions similaires (Dinstein et al., 2007; Chong et al., 2008; Kilner et al., 2009; Lingnau, Gesierich et Caramazza, 2009). De plus, ces effets ont lieu également lors de paradigmes cross-modaux, une suppression des réponses neuronales de ces régions étant observée lorsque les actions manuelles sont d'abord observées puis exécutées et vice versa (Kilner et al., 2009). Bien que débattu, l'effet de RS au sein de ce circuit fronto-pariétal observé aussi bien lors de l'observation que de l'exécution d'actions manuelles est largement discuté par rapport au rôle possible du système de neurones miroirs dans le codage des buts d'action.

Comme lors de la présente étude, bien que les aires pariétale inférieure et prémotrice soient les régions les plus consistantes à montrer un effet de RS, une diminution d'activité du lobule pariétal supérieur a également été observée dans de précédentes études lors de l'exécution répétée d'actions manuelles (Dinstein et al., 2007; Chong et al., 2008; Lingnau, Gesierich et Caramazza, 2009; Hamilton et Grafton, 2009). Pour les actions manuelles, il a été suggéré que le lobule pariétal supérieur soit impliqué dans des mécanismes de transformations visuomotrices et contienne des représentations d'action de bas niveau, alors que le lobule pariétal inférieur serait lui impliqué dans le codage des représentations de plus haut niveau du but de l'action (Jeannerod et Jacob, 2005). Dans ce cadre, puisqu'aucun feedback visuel n'a été bien sûr possible lors de l'exécution répétée des actions orofaciales, et du fait d'une contribution potentielle du lobule pariétal supérieur aux traitements visuospatiaux, une possible interprétation de l'effet de répétition suppression observé au niveau du précuneus gauche serait liée à des processus d'imagerie visuospatiale des mouvements orofaciaux. Bien que spéculative, cette hypothèse peut être reliée aux résultats d'une étude IRMf portant sur le traitement spatial des mouvements linguaux (Watanabe et al., 2004). Dans cette étude, les participants ont exécuté un mouvement lingual soit dirigé spécifiquement vers une dent (incisive, canine, molaire, supérieure ou inférieure, du côté gauche ou droit) soit consistant en une rétraction de la langue. Lors des mouvements dirigés vers une dent, une activité cérébrale accrue a été observée au sein des lobules pariétaux supérieur et inférieur et du cortex prémoteur. Ces

Grafton, Fagg et Arbib, 1998) de par leurs connections avec les régions préfrontales et pariétales. L'activation de ces régions a également été observée lors de mouvements orofaciaux simples (voir l'étude 1.1 du premier chapitre) et complexes (Dresel et al., 2005). L'effet RS dans la partie operculaire du gyrus frontal inférieur et au sein des aires visuelles telles le cortex extrastrié et le gyrus fusiforme pourrait vraisemblablement refléter les processus de perception des instructions visuelles liées à la tâche (Dehaene et al., 2001; Hamilton et Grafton, 2009). Néanmoins, puisque dans la présente étude les activations cérébrales dans ces régions n'ont pas survécu à une correction au niveau du cluster, leur interprétation reste limitée.

résultats pourraient partiellement découler d'une plus grande précision requise du codage du mouvement spatial de la langue dans le cadre d'une cible précise. De manière intéressante, une double dissociation entre imagerie motrice et visuelle et aires pariétales inférieure et supérieure a récemment été suggérée. Lors d'une étude utilisant la technique de stimulation magnétique transcrânienne (Pelgrims, Andres et Olivier, 2009), une perturbation temporaire de l'activité du lobule pariétal inférieur a en effet affecté sélectivement les temps de réaction lors d'une tâche d'imagerie motrice (rotation imaginée de la main) mais non lors d'une tâche d'imagerie visuelle (rotation de lettre) alors que les résultats inverses ont été observés suite à une perturbation de l'activité du lobule pariétal supérieur.

Pris ensemble, ces résultats ont des implications importantes sur les recherches liées au contrôle moteur, l'apprentissage sensorimoteur et la production de la parole. En effet, étant donné qu'il a été démontré lors de précédentes études que le sulcus intrapariétal et le lobule pariétal inférieur adjacent, le lobule pariétal supérieur et le cortex prémoteur ventral sont sensibles à l'effet de RS lors de l'exécution et/ou de l'observation répétée d'actions manuelles chez les humains et les primates non humains, nos résultats permettent dès lors d'étendre le rôle de ces régions liées au codage du but de l'action à des mécanismes adaptatifs/prédictifs aussi bien visuomoteurs que somatosensorimoteurs, dans le cas d'actions orofaciales intransitives où aucun feedback visuel n'est bien sûr présent.

Comme mentionné précédemment, alors que le codage du but de l'action et les processus de contrôle prédictif sensorimoteur ont été largement étudiés dans le cadre de mouvements des membres supérieurs, la plupart des modèles de production de la parole soutiennent également l'implication des cortex auditif et somatosensoriel dans les mécanismes de contrôle prédictif en ligne permettant d'évaluer les conséquences sensorielles prédites des actes de la parole avec les entrées sensorielles réelles, afin de contrôler/corriger par la suite la production (Guenther, 2006; Tian et Poeppel, 2010; Guenther et Vladusich, 2012; Hickok, Houde et Rong, 2011; Price, Crinion et MacSweeney, 2011). En appui de cette hypothèse, il a été démontré que des transformations transitoires et inattendues du feedback acoustique lors de la production de parole entraînent des ajustements articulatoires rapides en ligne (voir par exemple, Gracco et Löfqvist, 1994; Purcell et Munhall, 2006) et que ces adaptations persistent à court terme une fois une fois stoppée tout modification du feedback acoustique (par exemple, Houde et Jordan, 1998; Jones et Munhall, 2005). Au niveau neuronal, une modulation de l'activité du cortex auditif a été observée dans le cas d'un feedback acoustique altéré ou retardé (Hashimoto et Sakai, 2003; Heinks-Maldonado, Nagarajan et Houde, 2006; Christoffels, Formisano et Schiller, 2007; Tourville, Reilly et Guenther, 2008; Shiller et al., 2009). De même, bien que l'apport des informations somatosensorielles soit moins étudié dans le cadre de la production de la parole, des travaux comportementaux impliquant des perturbations orofaciales mécaniques de la peau du visage et des muscles du conduit vocal suggèrent également l'existence de mécanismes prédictifs/correctifs somatosensoriels lors de la production de parole (par exemple, Gracco et Abbs, 1985; Tremblay, Shiller et Ostry, 2003; Nasir et Ostry, 2006). Les résultats présents complètent ainsi ces travaux en démontrant le rôle des aires prémotrices ventrales et pariétales postérieures dans le codage du but d'action et les mécanismes prédictifs adaptatifs somatosensoriels lors de mouvements orofaciaux intransitifs et silencieux.

Limites de l'étude

Finalement, il est important de discuter certaines limites inhérentes à la présente étude. En premier lieu, les paramètres musculaires et cinématiques des mouvements orofaciaux n'ont pas été contrôlés lors de la session d'imagerie et, de fait, nous ne pouvons exclure la possibilité que l'effet observé de RS soit lié à une diminution de l'activité musculaire et/ou de vitesse et, de manière plus générale, à une fatigue musculaire engendrée par la répétition des mouvements orofaciaux. Face à cette possibilité, une diminution du signal BOLD a été observée dans le précuneus lors d'une étude IRMf portant sur la fatigue musculaire engendrée par des contractions maintenues et répétées des

muscles de la main (Van Duinen et al., 2007), un effet associé selon les auteurs à une pression artérielle accrue et non spécifiquement lié à l'activité neuronale (Critchley et al., 2000). Néanmoins, la fatigue musculaire implique classiquement une augmentation du signal BOLD dans les aires sensorimotrices (Liu et al., 2003; van Duinen et al., 2007), un effet probablement dû à des mécanismes de compensation permettant de maintenir une performance stable des actions produites. Enfin, il est également à noter qu'une variation des paramètres de force musculaire et cinématiques est fortement probable lors de la répétition des mouvements (Hamilton et Grafton, 2009). Sur ce point, les effets de RS observés ne pourraient donc pas dépendre de la réplication précise et identique d'une action d'un essai à l'autre mais bien du codage du but de cette action.

Une deuxième réserve méthodologique concerne le nombre limité d'essais par tâche motrice et par répétition. Pour cette raison, chacune des six occurrences a été modélisée indépendamment de l'articulateur et de la tâche motrice réalisée. Bien que nos résultats suggèrent que le codage du but des actions orofaciales et les mécanismes adaptatifs/prédictifs somatosensoriels sont codés dans les aires prémotrice et pariétales de l'hémisphère gauche, nous ne pouvons donc exclure que l'effet de RS observé soit dépendant d'un mouvement orofacial en particulier. Sur ce point, lors de l'analyse des réseaux neuronaux sous-jacents au contrôle moteur des gestes orofaciaux (voir également Partie 1.1), une activité plus importante a été observée pour les mouvements de rétraction de la langue dans une région du cortex prémoteur ventral proche de celle montrant un effet de RS. Des études complémentaires sont donc nécessaires afin de déterminer plus spécifiquement les régions sensibles à l'effet de RS pour chacun des mouvements de la langue, des lèvres et de la mandibule.

2.1.5 Conclusion

Cette étude IRMf utilisant un paradigme d'adaptation avait pour but de déterminer les corrélats neuronaux du codage du but des actions orofaciales et des mécanismes sensorimoteurs adaptatifs/prédictifs lors de la répétition de mouvements silencieux et intransitifs des lèvres, de la langue et de la mandibule. En accord avec de précédentes études portant sur des actions manuelles, une diminution de l'activité neurale lors de l'exécution répétée des actions orofaciales a été observée dans l'hémisphère gauche, au sein du sulcus intrapariétal et du lobule pariétal inférieur adjacent, du lobule pariétal supérieur et du cortex prémoteur ventral. Ces résultats suggèrent un codage du but des actions orofaciales intransitives et silencieuses au sein de ce circuit fronto-pariétal et démontrent l'importance de mécanismes adaptatifs/prédictifs motori-somatosensoriel dans le contrôle moteur orofacial.

Chapitre 2.2

Contrôle sensorimoteur adaptatif et but des actions vocaliques

Résumé / Abstract

Des études récentes portant sur les processus perceptifs et le contrôle moteur de la parole démontrent l'existence d'interactions sensorimotrices aussi bien lors de la production que de la perception de la parole. Lors de la production de parole, la modulation des réponses neuronales observées au sein des cortex auditif et somatosensoriel refléterait des mécanismes de contrôle des retours sensoriels, permettant la comparaison des conséquences sensorielles prédites des actes moteurs de parole et des entrées sensorielles réelles et ainsi un contrôle en ligne des productions de parole. Inversement, l'implication du système moteur observée lors de la perception de la parole pourrait correspondre à des mécanismes de simulation motrice dont le but serait de contraindre l'interprétation phonétique finale des entrées sensorielles par la génération interne de possibles candidats articulatoires. La présente étude en imagerie par résonance magnétique fonctionnelle, utilisant une méthode d'acquisition en sparse sampling et un paradigme de répétition suppression, avait pour but d'identifier les régions neurales impliquées dans des mécanismes d'adaptation sensorimotrice lors de la perception et de la production répétée de voyelles du Français. Ce paradigme de répétition suppression correspond à des processus d'adaptation et à une décroissance d'activité lors de l'observation ou exécution répétée d'actes moteurs et pourrait refléter un apprentissage sensorimoteur. Dans la présente étude, la perception et la production des voyelles ont activé un ensemble de régions sensorimotrices communes. De manière cruciale, une diminution de l'activité du gyrus frontal inférieur, du sulcus/gyrus temporal supérieur et moyen, du gyrus cingulaire antérieur et du lobule pariétal inférieur a été observée lors de la répétition des voyelles perçues et produites. Ces résultats suggèrent des mécanismes d'adaptation communs lors des traitements auditifs et des processus de contrôle moteur et démontrent l'existence d'un couplage fonctionnel entre systèmes de perception et de production de la parole.

Recent studies on speech processing and motor control provide evidence for sensorimotor interactions in both speech production and perception. During speech production, decreased neural responses observed within the auditory and somatosensory cortices are thought to reflect online feedback control mechanisms in which sensory consequence of the speech-motor act are evaluated with actual sensory input in order to further control production. Conversely, motor system activity observed during speech perception has been proposed to partly constrain phonetic interpretation of sensory inputs through the internal generation of candidate articulatory categorizations. We performed a sparse sampling functional magnetic resonance imaging study using a repetition-suppression (RS) paradigm to identify possible sensory-to-motor and motor-to-sensory loops involved in sensorimotor adaptation during repeated French vowel perception and production. The RS paradigm here refers to adaptation mechanisms and to decreased activity in specific neural populations due to repeated perceived or produced motor acts and has been proposed to reflect sensorimotor learning. In the present study, vowel perception and production activated a set of common sensorimotor brain areas. Crucially, activity in the inferior frontal gyrus, the posterior superior temporal sulcus and middle temporal gyrus, the anterior cingulate gyrus and the inferior

parietal lobule was suppressed during both vowel production and perception. These results strongly demonstrate common adaptive process during speech processing and motor control and provide evidence for a functional coupling of perception and production systems.

- *Mots clés : Perception et production de la parole, interactions sensorimotrices, codage du but de l'action, voie dorsale, fMRI, répétition suppression, adaptation, sparse sampling.*
- *Keywords: Speech perception and production, sensorimotor interactions, action goal coding, dorsal pathway, fMRI, repetition suppression, adaptation, sparse sampling.*

2.2.1 Introduction

L'étude présentée dans le Chapitre 1.2 portant sur les corrélats neuroanatomiques fonctionnels des systèmes de perception et de production des voyelles a permis de mettre en évidence ou tout du moins de suggérer : (1) un couplage fonctionnel de ces deux systèmes, notamment par l'observation d'activités motrices au sein du cortex prémoteur ventral et de la partie operculaire du gyrus frontal inférieur lors de la perception des voyelles et d'activités auditives au sein du gyrus supérieur temporal lors des deux tâches d'écoute et de production, (2) une implication de la voie dorsale et un possible codage distribué sensorimoteur des voyelles perçues et produites et, enfin, (3) une modulation de certaines régions motrices en fonction de l'expertise phonologique et articuloire du locuteur lors de la production des voyelles.

Néanmoins, face aux résultats de cette étude, la question d'un codage phonémo-spécifique des voyelles perçues et produites reste posée. En effet, étant donné les larges activations temporelles observées, des régions antérieures du gyrus/sulcus temporal supérieur jusqu'aux régions auditives postérieures et à la partie ventrale du gyrus supramarginal, nos résultats ne permettent ni d'appuyer ni de contredire l'existence d'une voie auditive antérieure (Scott et Johnsrude, 2003; Obleser et al., 2006; Rauschecker et Scott, 2009; Rauschecker, 2011) ou postérieure (Hickok et Poeppel, 2007) impliquée dans les processus de traitements acoustico-phonétiques. De même, nous n'avons pu mettre en évidence l'existence de régions topographiquement différenciées et spécifiquement impliquées dans le codage phonétique/phonémotopique des voyelles et des traits phonétiques sous-jacents.

Comme discuté précédemment, des analyses multivariées de reconnaissance statistique de patterns d'activation pourraient permettre de déterminer une possible ségrégation topographique entre les voyelles ou traits phonétiques perçus et produits (voir Formisano et al., 2008; Obleser et al., 2010)⁴⁹. De manière plus générale, l'existence de traitements spécifiques liés au codage phonétique et phonémotopique des voyelles renvoie également à la question de régions neurales dédiées au codage du but des actions vocaliques perçues ou produites. Face à cette question, l'étude précédente (voir Chapitre 2.1) a permis d'étendre le rôle de régions spécifiques liées au codage du but de l'action - soit le sulcus intrapariétal et le lobule pariétal inférieur adjacent, le lobule pariétal supérieur et le cortex prémoteur ventral de l'hémisphère gauche - à des mécanismes adaptatifs/prédictifs somatosensorimoteurs lors de la production répétée d'actions orofaciales intransitives. La présente étude a pour but d'étendre ces résultats au codage du but des actions vocaliques perçues ou produites par l'utilisation d'un paradigme de répétition suppression. L'utilisation d'un tel paradigme a ainsi pour but de déterminer les régions neurales sensibles aux mécanismes sensorimoteurs adaptatifs mis en œuvre lors de production et de l'écoute répétée d'une même voyelle. La mise en évidence de telles régions impliquées dans le codage du but des actions vocaliques apparaît également importante face à une possible organisation hiérarchique des représentations de parole au sein du système nerveux central et en regard de certains modèles

⁴⁹ De telles analyses sont actuellement en cours sur nos données dans le cadre d'une collaboration avec Michel Dojat et Mathieu Ruiz de l'Institut des Neurosciences de Grenoble.

neurobiologiques de production et de perception de la parole postulant l'existence de boucles prédictives/correctives (Scott et Johnsrude, 2003; Callan et al., 2004; Guenther, 2006; Wilson et Iacoboni, 2006; Skipper et al., 2007; Poeppel, Idsardi et van Wassenhove, 2008; Rauschecker et Scott, 2009; Tian et Poeppel, 2010; Hickok, Houde et Rong, 2011; Price, Crinion et MacSweeney, 2011; Rauschecker, 2011; Guenther et Vladusich, 2012).

Codage du but des actions vocaliques lors de la perception et de la production de parole

Comme mentionné précédemment (voir Chapitre 2.1), des enregistrements neurophysiologiques unicellulaires effectués chez le primate non humain et des études IRMf utilisant un protocole d'adaptation chez l'homme ont permis de préciser le rôle de régions pariétales postérieures et prémotrices ventrales dans le codage du but des actions manuelles (Rizzolatti et al., 1988; Fogassi et al., 2005; Dinstein et al., 2007; Gardner et al., 2007; Chong et al., 2008; Rozzi et al., 2008; Umiltà et al., 2008; Kilner et al., 2009; Hamilton et Grafton, 2009; Bonini et al., 2010, 2011). De manière importante, ce codage du but des actions a été mis en évidence aussi bien lors de l'observation que lors de la production d'actions manuelles et a été largement discuté en regard du système de neurones miroirs. Face à ces résultats et à ceux observés dans les études présentées dans les Chapitres 1.2 et 2.1, nous avons émis l'hypothèse d'un effet d'adaptation phonémo-spécifique au sein des régions de la voie dorsale codant de manière sélective les cibles phonémiques/vocaliques (acoustiques, articulatoires ou sensorimotrices), soit le cortex prémoteur ventral et la partie operculaire du gyrus frontal inférieur, les régions temporales antérieures et/ou postérieures et le lobule pariétal inférieur et le sillon pariétal de l'hémisphère gauche.

Il convient cependant de rappeler les résultats observés lors de précédentes études IRMf qui ont exploité le paradigme de répétition suppression dans le but de déterminer les régions neuronales sensibles à des mécanismes d'adaptation lors de la perception et production de la parole et du langage.

En ce qui concerne la perception de la parole ou du langage, Hasson, Nusbaum et Small (2006) ont étudié les effets possibles d'adaptation au sein de régions impliquées dans les traitements sémantiques et syntaxiques lors de l'écoute passive ou lors d'une tâche de compréhension de phrases grammaticalement correctes ou non. Les deux tâches ont induit un effet de RS dans les régions temporales bilatérales (incluant les gyri temporaux supérieur, transverse et moyen), alors que la tâche de jugement grammatical impliquait de plus des effets d'adaptation au niveau du gyrus frontal inférieur (BA 44/47) et du cortex insulaire. Les auteurs concluent de ces résultats que l'effet de RS observé lors des deux tâches dans les régions auditives refléterait une meilleure compréhension des phrases perçues (aussi bien du point de vue phonologique, sémantique que syntaxique) tandis que celui observé spécifiquement lors de la tâche de jugement dans le gyrus frontal inférieur serait relié à une meilleure réalisation de la tâche par les participants. Face à cette dernière hypothèse, les auteurs ne discutent cependant pas d'une différence d'implication des traitements phonologiques, sémantiques et/ou syntaxiques entre ces deux tâches passive et active et, de là, la possibilité que l'effet de RS observé dans le gyrus frontal inférieur puisse justement découler des ces différences.

Vaden, Muftuler et Hickok (2010) se sont intéressés à un possible effet de RS fonction de ce que les auteurs nomment un "effet de répétition phonologique", à savoir la répétition de structures CV initiales et/ou VC finales lors de l'écoute d'une liste de 4 mots de type CVC (consonne-voyelle-consonne). Trois conditions étaient contrastées consistant en l'écoute 1) de 4 mots ne partageant aucune structure CV ou VC (par exemple, "jug", "knit", "rage", "hem"), 2) de 4 mots incluant la même structure CV initiale (par exemple, "cab", "calf", "cat", "cap") et 3) de 4 mots identiques. Une analyse paramétrique reliée à ces 3 conditions a montré une activité bilatérale au sein de la partie moyenne du sulcus temporal supérieur inversement fonction de l'effet de répétition phonologique (soit une

activité moindre dans le cas de l'écoute de 4 mots identiques), en accord avec le rôle possible selon les auteurs de cette région dans les traitements phonologique et lexicaux (Hickok et Poeppel, 2007). Bien que peu discutées⁵⁰, d'autres régions montrent cependant un effet de RS fonction de l'effet de répétition phonologique, l'aire motrice supplémentaire et le gyrus cingulaire, le gyrus frontal inférieur et le gyrus frontal moyen, le gyrus précentral et le précunéus.

D'autres études ont utilisé le paradigme de RS afin mettre en évidence les structures spécifiquement impliquées dans la perception intra- et inter-catégorielle des sons de parole, en présentant des stimuli issus d'un continuum /ga/-/da/ (Joanisse, Zevin et McCandliss, 2007) ou /ba/-/da/ (Raizada et Poldrack, 2007). Joanisse et al. (2007) montrent des activations plus importantes au niveau du sulcus temporal supérieur dans sa partie moyenne, du gyrus temporal moyen et du lobule pariétal inférieur de l'hémisphère gauche lors de l'écoute répétée de deux stimuli appartenant à deux catégories phonologiques différentes, par rapport à deux stimuli appartenant à la même catégorie phonologique. Pour les auteurs, ce résultat suggère le rôle de ces régions dans le codage des traits phonétiques des syllabes perçues. De même, Raizada et Poldrack (2007) montrent une amplification sélective des réponses neuronales au niveau du gyrus supramarginal gauche (et de façon moindre dans le cervelet droit, le gyrus cingulaire et le cortex préfrontal) lors de l'écoute de deux stimuli n'appartenant pas à la même catégorie syllabique, mais pas si la même différence acoustique a lieu au sein de la même catégorie phonologique. L'effet d'amplification sélective observé, inversement relié à un effet de RS⁵¹, indique selon les auteurs le rôle du gyrus supramarginal gauche dans les processus inhérents à la perception catégorielle de la parole. De manière importante, un tel effet n'a pu être mis en évidence au sein de régions auditives (gyrus de Heschl et planum temporale).

Enfin, Hasson et al. (2007) ont utilisé le paradigme d'adaptation pour mettre en évidence les régions cérébrales non directement sensibles aux entrées sensorielles (auditives et visuelles) mais à l'interprétation perceptive du stimulus. Pour ce faire, ils ont utilisé l'effet McGurk qui consiste généralement à présenter auditivement une syllabe contenant une consonne initiale bilabiale (par exemple /pa/) et visuellement une syllabe contenant une consonne initiale vélaire (par exemple /ka/) et pouvant induire un effet d'illusion perceptive (fusion audiovisuelle) où la syllabe perçue ne correspond ni au stimulus auditif ni au stimulus visuel (par exemple /ta/). Dans cette étude, une cible audiovisuelle incongruente (/pa/-audio, /ka/-visuel) était précédée a) d'un stimulus audiovisuel congruent partageant les mêmes informations auditives (/pa/), b) les mêmes informations visuelles (/ka/) ou c) perçue comme la syllabe audiovisuelle incongruente (/ta/). Une moindre activité au sein de la partie operculaire du gyrus frontal inférieur et du planum polaire de l'hémisphère gauche a été observée dans le dernier cas, lorsque les deux syllabes ne partageaient aucune information auditive et visuelle mais dont l'interprétation perceptive était identique. Pour les auteurs, ce résultat indique que ces régions coderaient les représentations de parole à un niveau abstrait indépendamment de leurs propriétés acoustiques ou visuelles.

En ce qui concerne la production de parole, Peeva et al. (2010) ont étudié les régions cérébrales sensibles à l'effet d'adaptation en fonction de la répétition de phonèmes, de syllabes ou de séquences bisyllabiques. Pour ce faire, quatre conditions étaient contrastées consistant en six productions a) d'une même logatome bisyllabique (par exemple, "flu.blo" x6), b) de deux logatomes dont les syllabes étaient inversées (par exemple, "ze.klo", "klo.ze" x3), c) de deux logatomes différents mais partageant les mêmes phonèmes (par exemple, "li.muf" "mu.fli" x3) ou d) de 6 logatomes différents (par exemple, "gu.pri", "dra.vo", "vi.bag", "nu.vaf", "bu.lif", "bas.mo"). Ainsi, les régions sensibles à l'effet d'adaptation phonémique devraient montrer une diminution d'activité plus marquée dans les conditions a, b et c, celles sensibles à l'effet d'adaptation syllabique dans les conditions a et b et, enfin, celles sensibles à l'effet d'adaptation supra-syllabique dans la condition a.

⁵⁰ Les auteurs indiquent simplement, qu'en dehors du sulcus temporal, les autres régions sensibles à l'effet de RS pourraient jouer un rôle dans des mécanismes attentionnels et/ou de codage prédictif.

⁵¹ Cet effet d'amplification sélective est calculé en contrastant l'activité observée lors de la présentation de paires de stimuli S1 et S2 différents ou identiques selon l'équation : (S1 puis S2 + S2 puis S1) – (S1 puis S1 + S2 puis S2).

Un effet d'adaptation phonémique a ainsi été trouvé dans l'aire motrice supplémentaire, les ganglions de la base (pallidum), la partie postérieure du gyrus supérieur temporal, et le cervelet supérieur. Les régions sensibles à l'effet d'adaptation syllabique ont été trouvées dans le cortex prémoteur ventral gauche et celles sensibles à un effet d'adaptation supra-syllabique dans le cervelet supérieur droit. A noter que le cortex moteur primaire ainsi que les aires auditives primaires n'ont montré de sensibilité à aucun de ces effets d'adaptation.

Cadre de l'étude

En regard de ces travaux, la présente étude IRMf avait pour but de déterminer les régions neurales sensibles à un effet de répétition suppression lors de l'écoute ou de la production répétée d'une même voyelle. Comme indiqué précédemment, nous avons émis l'hypothèse d'un effet d'adaptation phonémo-spécifique au sein des régions de la voie dorsale codant de manière sélective les cibles vocaliques, soit le cortex prémoteur ventral et la partie operculaire du gyrus frontal inférieur, les régions temporales antérieures et/ou postérieures et le lobule pariétal inférieur et le sillon pariétal de l'hémisphère gauche. De manière à caractériser un possible effet d'adaptation, la présentation ou production des voyelles était réalisé par trains consécutifs de 6 voyelles identiques. A l'exception de l'utilisation du paradigme d'adaptation, le protocole expérimental était en tous points identiques à celui utilisé lors de l'étude présentée dans le Chapitre 1.2. Les participants ont d'abord écouté passivement 9 voyelles différentes maintenues du français (multiples occurrences des voyelles /i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/), propres à chaque sujet et enregistrées préalablement à l'étude. Dans la tâche suivante, réalisée lors de la même session IRM et en utilisant exactement les mêmes paramètres d'acquisition, les participants ont produit ouvertement ces mêmes voyelles. Egalement, la technique d'acquisition en sparse sampling a été utilisée dans le but de réduire le bruit du scanner lors des tâches de perception et de production des voyelles ainsi que d'éliminer de possibles artefacts liés aux mouvements articulatoires lors de la tâche de production des voyelles.

2.2.2 Méthode

Protocole

Participants

Douze volontaires adultes sains ont participé à l'étude (8 hommes; âge : 20-42 ans, moyenne : 29 ans). Le recrutement des sujets est explicité dans la partie "Considérations Méthodologiques" du premier chapitre.

Stimuli

Comme pour la précédente étude (voir Chapitre 1.2), un premier entretien préalable à la session IRM avait pour but d'expliquer à chaque participant le déroulement général de l'étude, les différentes tâches perceptives et motrices ainsi que de procéder à un enregistrement des voyelles dans une chambre sourde en vue de les utiliser pour la tâche de perception de parole.

Chaque voyelle a été produite 10 fois selon un ordre pseudo-aléatoire (de manière à ce qu'une même voyelle ou un même trait phonétique d'arrondissement, d'antériorité ou d'ouverture ne soit pas jamais répété). Pour chaque essai, une consigne visuelle ("ii", "ou", "ée"...) indiquait au sujet la voyelle à produire. Précédée d'une croix de fixation durant 500ms, l'affichage de cette consigne durant 1s servait également de repère temporel aux sujets pour la réalisation et la production maintenue de la voyelle.

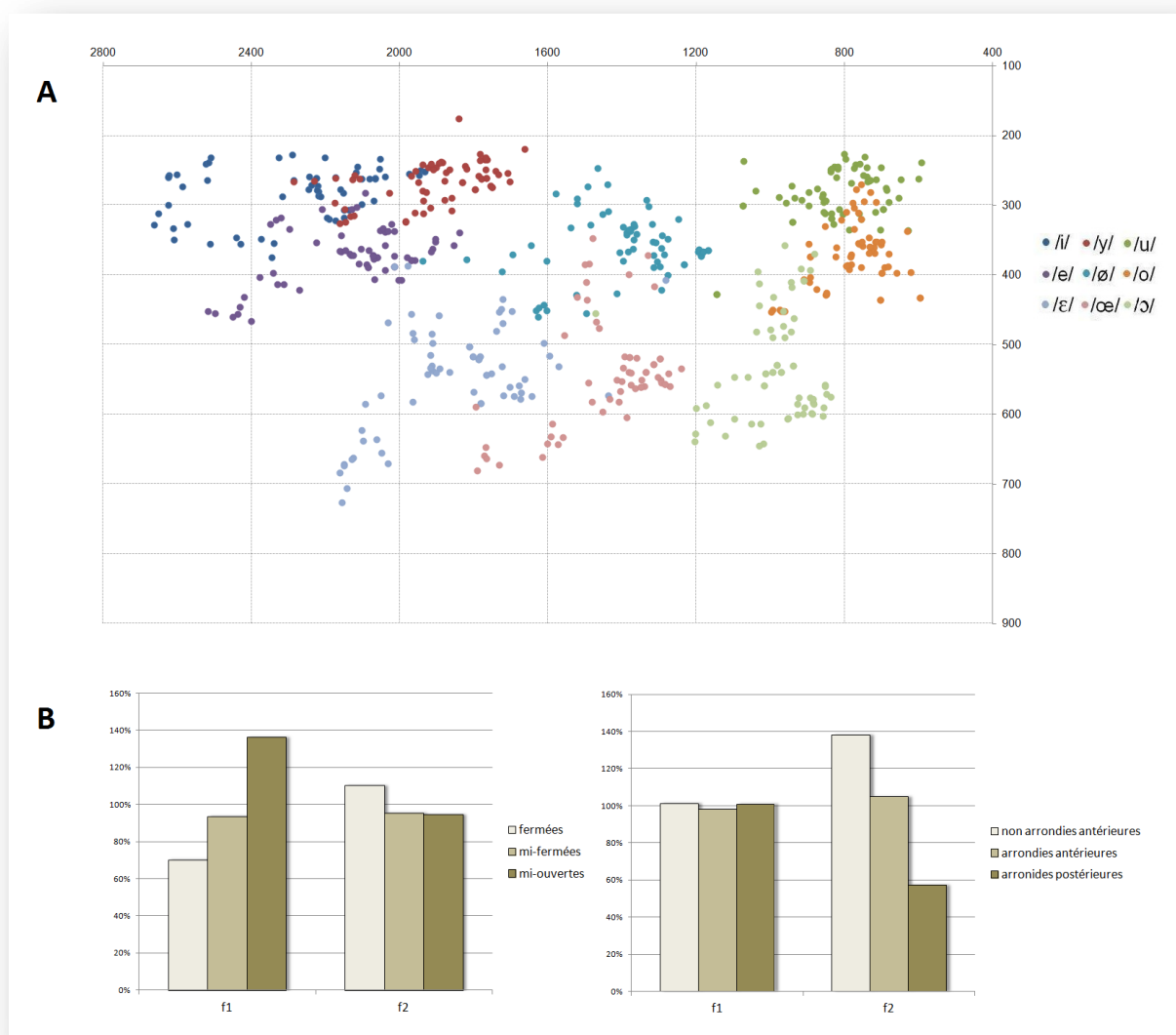


Figure 2.8 : A) Espace acoustique F_1/F_2 des 9 voyelles du Français enregistrées préalablement à l'étude IRM pour chacun des sujets. Chaque point représente la valeur moyenne de six occurrences d'une même voyelle pour un seul sujet. Les valeurs moyennes pour chaque type de voyelle pour les 12 sujets sont également indiquées (les barres d'erreurs représentent l'erreur standard; voir Table 2.3). B) Valeurs formantiques moyennes F_1/F_2 des voyelles exprimées en pourcentage de la moyenne générale pour les voyelles fermées, mi-fermées et mi-ouvertes (gauche) et pour les voyelles antérieures non arrondies, antérieures arrondies et postérieures arrondies (droite).

Voyelle	Ouverture	Antériorité	Arrondissement	F_0	F_1	F_2	Durée
/i/	Fermée	Antérieure	Non arrondie	156 ±15	279 ±11	2260 ±67	0,672 ±0,050
/y/	Fermée	Antérieure	Arrondie	153 ±15	269 ±9	1945 ±53	0,672 ±0,056
/u/	Fermée	Postérieure	Arrondie	156 ±16	284 ±10	815 ±23	0,683 ±0,053
/e/	Mi-fermée	Antérieure	Non arrondie	150 ±14	376 ±12	2169 ±58	0,666 ±0,052
/ø/	Mi-fermée	Antérieure	Arrondie	151 ±14	363 ±14	1443 ±51	0,682 ±0,059
/o/	Mi-fermée	Postérieure	Arrondie	153 ±15	373 ±14	790 ±25	0,679 ±0,057
/ɛ/	Mi-ouverte	Antérieure	Non arrondie	151 ±14	552 ±23	1893 ±59	0,676 ±0,064
/œ/	Mi-ouverte	Antérieure	Arrondie	150 ±14	538 ±24	1473 ±49	0,650 ±0,045
/ɔ/	Mi-ouverte	Postérieure	Arrondie	149 ±14	538 ±24	1001 ±31	0,674 ±0,045

Table 2.3 : Valeurs moyennes de la fréquence fondamentale (F_0 en Hz), des premier et second formants (F_1 , F_2 en Hz) et de la durée (en ms) pour toutes les voyelles préenregistrées auprès des participants et utilisées comme stimuli dans la tâche de perception. Les erreurs standards sont indiquées.

Il était demandé aux participants d'initier chacune de ces tâches depuis une position immobile de repos, les muscles linguaux et labiaux relâchés et de revenir à cette position initiale de repos à la fin du mouvement. L'intervalle entre chaque essai était de 5s. Les paramètres d'enregistrement des voyelles étant similaires à ceux utilisées lors de la session IRM, cette procédure servait donc également d'entraînement à la tâche de production dans le scanner.

En tout, 1080 voyelles (12 sujets x 9 voyelles x 10 productions) ont été enregistrées indépendamment avec un taux d'échantillonnage de 44.1 kHz et une résolution de 16 bits. Pour chaque voyelle et chaque participant, la fréquence fondamentale (F_0), le premier et second formant (F_1 et F_2) et la durée de production ont été calculés à l'aide du logiciel Praat (Institute of Phonetic Sciences, University of Amsterdam, NL; voir Figure 2.8 et Table 2.3). De manière attendue (Schwartz et al., 1997a,b; Ladefoged, 2006; Ménard et al., 2008), les valeurs moyennes de F_1 des voyelles fermées, mi-fermées et mi-ouvertes étaient fonction du degré d'ouverture de la voyelle (en moyenne, 281Hz, 373Hz et 537Hz, respectivement). Parallèlement, les valeurs moyennes de F_2 des voyelles non arrondies antérieures, arrondies antérieures et arrondies postérieures étaient positivement corrélées avec le degré de (non)arrondissement et d'antériorité (en moyenne, 2050Hz, 1584Hz et 1088Hz, respectivement).

Procédure expérimentale

La passation IRM consistait en 2 tâches d'écoute passive et de production ouverte des voyelles /i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/. En vue de minimiser d'éventuels effets de simulation motrice, la tâche de perception auditive a toujours précédé la tâche de production.

Perception. La tâche de perception a été réalisée lors des trois premiers runs fonctionnels (voir partie "Considérations Méthodologiques", Figure 1.1). Elle consistait en l'écoute passive d'une des 9 voyelles maintenues du Français (/i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/) propres à chaque sujet et enregistrées préalablement à l'étude (6 occurrences différentes par voyelle ont été présentées). Une condition de repos, sans aucun mouvement ni stimulation auditive, servait de condition contrôle. Durant l'expérience, l'écoute d'une voyelle ou une condition de repos apparaissait toutes les 10 secondes (du fait de l'acquisition des images fonctionnelles en sparse sampling) par série de 6 occurrences consécutives d'un type de voyelle. Chaque répétition de voyelle (1 à 6) consistait en 27 essais (3 x 9 voyelles distinctes). Concernant le protocole d'acquisition en sparse sampling, l'intervalle de temps séparant la perception d'une voyelle et le milieu d'acquisition de l'image fonctionnelle correspondante était contrebalancé entre chaque essai et était de 4s, 5s ou 6s. 180 scans fonctionnels ont ainsi été acquis ((6 répétitions * 27 essais) + (1 condition de repos x 18 scans) ou encore (9 voyelles + 1 condition de repos) * 18 scans). Lors de chaque run fonctionnel, 3 scans ont été ajoutés en début de session pour équilibrer le signal IRM et ont ensuite été supprimés des analyses.

Production. La tâche de production était réalisée lors des 3 runs fonctionnels suivants (voir partie "Considérations Méthodologiques", Figure 1.1). Il était demandé aux participants de produire ouvertement chacune des 9 voyelles de manière maintenue. Tout comme lors de la tâche de perception, une condition contrôle, sans mouvement ni stimulation auditive, a été ajoutée. Le sujet devait produire l'une des 9 voyelles ou la condition de repos toutes les 10 secondes par série de 6 occurrences identiques consécutives. Chaque répétition de voyelle (1 à 6) consistait en 27 essais (3 x 9 voyelles distinctes). Comme pour la tâche de perception, l'intervalle de temps séparant la production d'une voyelle et le milieu d'acquisition de l'image fonctionnelle correspondante était contrebalancé entre chaque essai et était de 4s, 5s ou 6s. Pour chaque essai, une consigne visuelle ("ii", "ou", "repos"...) indiquait au sujet la condition à réaliser. Précédée d'une croix de fixation durant 500ms, l'affichage de cette consigne durant 1s servait également de repère temporel aux sujets pour la réalisation de la tâche. Il était demandé aux participants d'initier chacune de ces tâches depuis une position immobile de repos, les muscles linguaux et labiaux relâchés et de revenir à cette position initiale de repos à la fin du mouvement. Un enregistrement des productions du sujet a été

réalisé de manière à contrôler *a posteriori* d'éventuelles erreurs de production et de les supprimer des analyses suivantes. 180 scans fonctionnels ont ainsi été acquis ((6 répétitions * 27 essais) + (1 condition de repos x 18 scans) ou encore (9 voyelles + 1 condition de repos) * 18 scans). Lors de chaque run fonctionnel, 3 scans ont été ajoutés en début de session pour équilibrer le signal IRM et ont ensuite été supprimés des analyses.

Prétraitements

Préalablement aux analyses statistiques, une série de prétraitements des images fonctionnelles a été réalisée avec le logiciel SPM 5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) sous environnement Matlab 7.1 (Mathworks, Natick, MA). Ces prétraitements incluaient des procédures de réalignement, de normalisation dans l'espace commun du Montreal Neurological Institute (repère MNI) et de lissage spatial (voir partie "Considérations Méthodologiques" pour une description détaillée des ces procédures).

Erreurs de production. L'écoute et l'inspection du signal acoustique de tous les essais (incluant 1944 voyelles produites et 216 essais de repos) ont été effectuées pour s'assurer de l'exactitude des réponses des participants. Trois types d'erreurs ont été observés : des omissions, des productions erronées ou des hésitations. En tout, 17 images fonctionnelles relatives à ces erreurs de productions ont été écartées des analyses (en moyenne, 0.73% \pm 0.15%). Cette analyse démontre la réalisation correcte de plus de 99% des voyelles produites (aucun participant n'ayant commis plus de 10% d'erreurs).

Analyses statistiques

Suite aux prétraitements, les données fonctionnelles ont été analysées avec le logiciel SPM 5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) sous environnement Matlab 7.1 (Mathworks, Natick, MA). Les régions cérébrales activées ont été labellisées en utilisant la toolbox SPM Anatomy (Eickhoff et al., 2005) et, si nécessaire, en utilisant le logiciel Talairach Daemon (Lancaster et al., 2000). Les analyses par région d'intérêt (ROI) ont été effectuées via la toolbox SPM Anatomy. Pour la visualisation des résultats, les cartes d'activation ont été superposées sur un template anatomique standard en utilisant le logiciel MRICRON (<http://www.sph.sc.edu/comd/rorden/mricron/>).

Analyses individuelles

Au niveau individuel et suite aux prétraitements des images fonctionnelles (voir partie "Considérations Méthodologiques"), les activations neuronales liées à la perception et production des 9 voyelles et traits phonétiques correspondants ont été estimées par l'utilisation d'un Modèle Linéaire Général (GLM; Friston et al., 1995). De manière à déterminer un possible effet de RS, un second GLM a été estimé pour chaque participant afin d'évaluer les activités reliées aux 6 répétitions successives lors des tâches de perception et de production.

Les modèles incluaient des régresseurs d'intérêt reliés soit aux 2 tâches perceptives et motrices pour chacune des 9 voyelles (2 x 9 régresseurs représentés chacun par 18 images fonctionnelles) indépendamment des 6 répétitions, soit aux 6 répétitions pour les deux tâches motrices et perceptives (2 x 6 régresseurs représentés chacun par 27 images fonctionnelles) et des régresseurs de non-intérêt liés aux paramètres de réalignement. Les conditions de repos formaient une ligne de base (2 x 18 images fonctionnelles).

Pour les deux modèles, la réponse de type hémodynamique associée à chaque événement a été modélisée par une fonction à réponse impulsionnelle finie et à impulsion unique (FIR) pour chaque

scan fonctionnel. Avant l'estimation du modèle, un filtrage des basses fréquences *a priori* non-relées aux conditions expérimentales (variations lentes d'origine physiologique) a été appliqué (passe-haut de fréquence de coupure de 1/128 Hz). Les estimations des réponses FIR modélisées ont été calculées de manière à correspondre avec le déroulement temporel du signal BOLD observé pour chaque voxel et chaque condition. Pour chaque participant, des cartes d'analyse statistique individuelles ont ainsi été calculées pour chaque voyelle perçue ou produite dans le premier modèle et pour chacune des 6 répétitions pour les deux tâches motrices et perceptives dans le second modèle, par contraste avec la condition contrôle. Ces cartes statistiques individuelles ont été ensuite utilisées pour les analyses de groupe.

Analyses de groupe

Afin de permettre une généralisation des résultats observés pour la population étudiée (Friston, Holmes et Worsley, 1999), des analyses de groupe de second niveau à effet aléatoire ont été réalisées au moyen d'analyses de variance (ANOVA) à mesures répétées.

Perception et production des voyelles

De manière à déterminer les activités cérébrales liées à la perception et la production de voyelles, une ANOVA à mesures répétées a été réalisée avec comme facteurs intra-sujet la tâche (2 niveaux, perception ou production) et la voyelle perçue ou produite (9 niveaux relatifs à chacune des voyelles). La variable "sujets" a été prise en compte comme facteur aléatoire.

Réseaux généraux : Deux contrastes t unidirectionnels ont été calculés afin de déterminer les régions cérébrales activées de manière spécifique lors de la perception et lors de la production des voyelles (toutes voyelles confondues) par rapport à la condition de repos. Dans le but d'identifier les activations communes/partagées aux 2 tâches, une analyse de conjonction (Friston et al., 1999; Nichols et al., 2005) a également été effectuée.

Traits phonétiques : Séparément pour les tâches de perception et de production, des analyses supplémentaires ont été réalisées en regard du trait d'ouverture, d'une part, et des traits d'arrondissement et d'antériorité, d'autre part. Ainsi, aussi bien pour la tâche de perception que pour celle de production, des contrastes t ont été calculés pour les voyelles fermées, mi-fermées et mi-ouvertes et pour les voyelles antérieures non arrondies, antérieures arrondies et postérieures arrondies. Enfin, des contrastes F bidirectionnels correspondant à l'effet principal du trait d'ouverture et du trait d'arrondissement/antériorité ont été calculés dans le but de déterminer les différences d'activations cérébrales entre types de voyelles.

Effet de RS

Une seconde analyse de groupe a été réalisée dans le but de comparer les activations cérébrales fonction de l'ordre de répétition des voyelles lors des tâches de perception et de production, indépendamment du type de voyelle perçue ou produite. Pour ce faire, une seconde ANOVA à mesures répétées a été effectuée avec comme facteurs intra-sujet la tâche (perception ou production) et l'ordre de répétition (incluant 6 niveaux). La variable "sujets" a été prise en compte comme facteur aléatoire.

Pour chaque tâche, six contrastes t ont été calculés de manière à déterminer les régions cérébrales activées de manière spécifique pour chacune des répétitions par rapport à la ligne de base. Dans le but d'identifier un effet de RS, un contraste t modélisant une décroissance linéaire d'activité au cours de ces 6 répétitions a également été calculé pour les deux tâches. Enfin, une analyse de conjonction des effets de RS observés lors de la perception et de la production de voyelles a été estimée.

2.2.3 Résultats

Les résultats des analyses de groupe sont reportés selon un seuil statistique défini à $p < .05$ après correction de type “False Discovery Rate” (FDR; Genovese et al., 2002), excepté pour les effets principaux des traits phonétiques et l'analyse de conjonction des effets de RS reportés à un seuil statistique de $p < .001$ non corrigé, et une taille minimale des clusters de 25 voxels. Pour toutes les analyses et contrastes, les pics d'activation maximum ont été déterminés dans chaque cluster, leur localisation a ensuite été labellisée avec la boîte à outils Anatomy de SPM (Eickhoff et al., 2005). Si une région n'avait pas pu être assignée avec Anatomy, elle a été déterminée avec le logiciel Talairach Daemon (Lancaster et al., 2000) grâce aux coordonnées du pic d'activation converties de l'espace MNI à l'espace stéréotaxique standard de Talairach et Tournoux (1988).

Perception et production des voyelles

La Table 2.4 répertorie les activations principales pour les tâches de perception et de production (par rapport à la ligne de base) ainsi que les activations communes révélées par l'analyse de conjonction. Les cartes d'activations sont illustrées dans la Figure 2.9.

Perception

La perception de voyelles a induit des activations bilatérales des gyri/sulci temporaux supérieurs, incluant aussi bien les aires auditives primaires au niveau du gyrus de Heschl (BA 41), que les aires auditives secondaires et associatives (BA 22). Par ailleurs, on trouve une activation de l'aire temporopolaire droite (BA 38). Des activations bilatérales ont été observées au niveau du cortex insulaire ainsi que des activations à gauche au niveau du gyrus frontal inférieur dans sa partie operculaire (BA 44) et dans le gyrus supramarginal (BA 40).

Production

La production de voyelles a induit de larges activations bilatérales : depuis l'aire motrice supplémentaire s'étendant aux cortex prémoteur, primaire sensorimoteur, les gyri frontaux moyens (BA 8, 9 à gauche et BA 46 à droite) et inférieurs (BA 44, 45 à droite et BA 47 à gauche), les aires postérieures pariétales y compris l'aire somatosensorielle primaire (BA 3) et associative (BA 5), l'opercule rolandique (BA 43) le précuneus (BA 7) et le gyrus supramarginal (BA 40). Des activations ont été également observées au niveau du système limbique : les aires cingulaires antérieure (partie ventrale, BA 24 et dorsale, BA 32) et postérieure (partie dorsale, BA 31 et rétrospléniale, BA 29, 30), le thalamus (y compris le noyau latéral postérieur ventral et le corps génouillé médian), le noyau caudé (tête et corps), ainsi qu'au niveau du putamen, le pallidus, le noyau rouge et la substance noire. D'autres activations bilatérales ont été observées dans le culmen et le déclive du cervelet et dans les cortex striés et extrastriés du lobe occipital. Des activations ont également été reportées dans les gyri/sulci temporaux supérieurs (BA 41, 22) et inférieurs (BA 37) de manière bilatérale ainsi que dans l'insula droite.

Conjonction

L'analyse de conjonction des tâches de perception et de production a révélé des activations communes au niveau du gyrus/sulcus temporal supérieur (incluant les cortex primaire, secondaire et associatif).

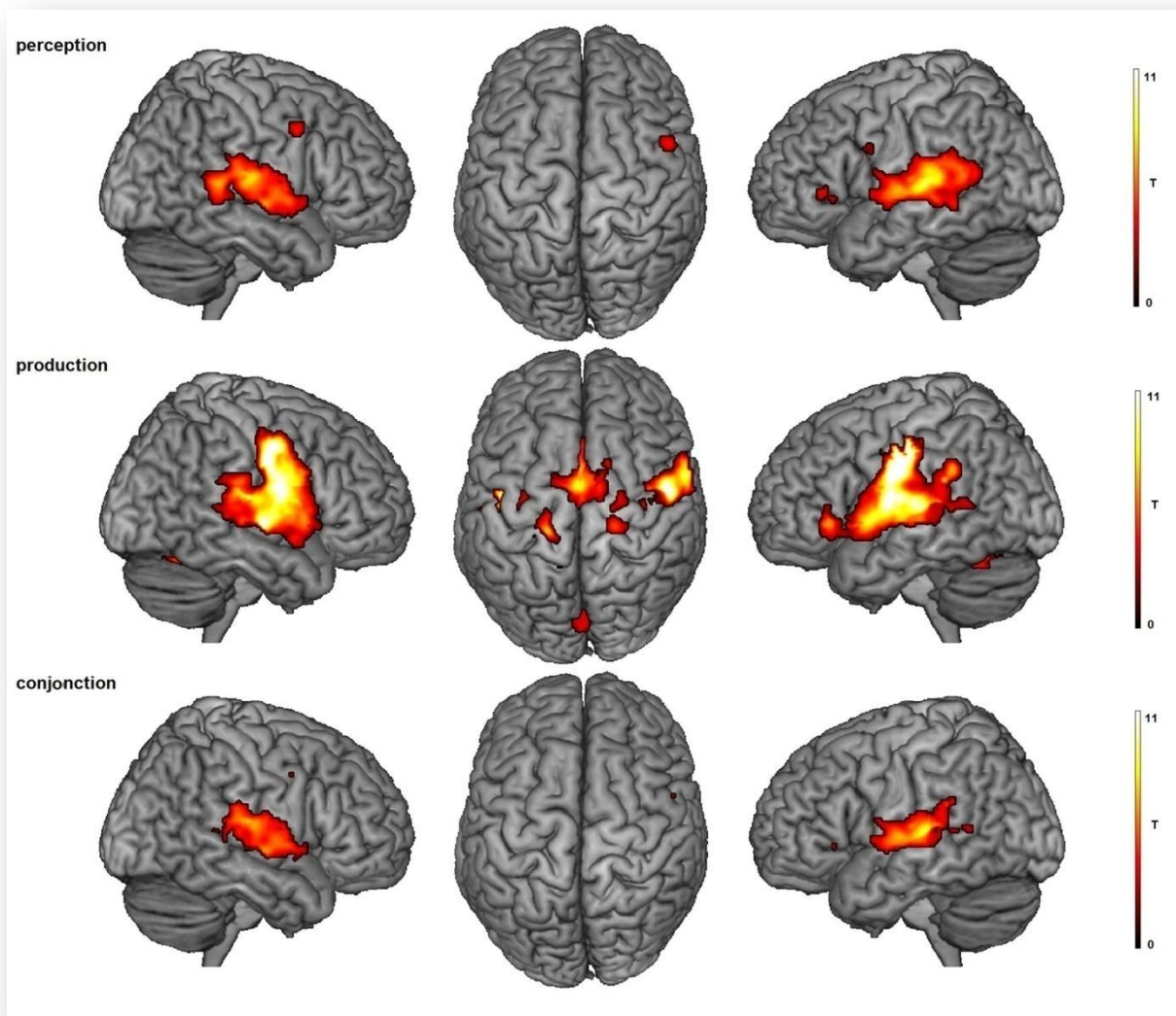


Figure 2.9 : Vue surfacique des régions cérébrales spécifiquement activées lors des tâches de perception et de production des voyelles et des activations communes aux deux tâches révélées par l'analyse de conjonction. Résultats reportés à un seuil corrigé (FDR, $p < .05$) et une taille minimale des clusters de 25 voxels.

Régions	H	BA	Perception				BA	Production				BA	Conjonction			
			x	y	z	T		x	y	z	T		x	y	z	T
Frontales																
Gyrus frontal inférieur	L	44	-42	6	8	3.82										
	R						44	58	8	22	5.97					
	R						45	54	34	6	5.06					
Cortex prémoteur	L						47	-36	36	4	4.02					
	L						6	-20	-20	56	4.71					
	R						6	50	-6	42	7.78					
Cortex moteur primaire	R						6	60	2	22	5.93					
	L						4	-44	-14	36	10.49					
	L						4	-20	-30	56	6.28					
Gyrus frontal moyen /	R						4	46	-8	32	8.56					
	R						4	20	-28	58	3.70					
	L						8	-14	32	42	4.62					
Cortex préfrontal	L						9	-32	18	38	4.99					
Gyrus frontal médial /	L						6	-20	6	54	4.00					
Aire motrice supplémentaire	L						6	-2	0	52	3.53					

<u>Temporales</u>																
Aire temporopolaire	R	38	46	10	-12	3.92										
Gyrus temporal supérieur	L	22	-46	-22	4	5.32	22	-50	8	-2	6.21					
	L						22	-52	-36	8	5.65					
	R	22	60	-6	6	4.90	22	52	-6	4	6.28	22	58	-6	6	4.72
	R	22	54	-38	2	4.85										
	R	22	48	-2	-8	4.69										
Gyrus temporal transverse	L	41	-56	-28	8	6.00	41	-44	-30	12	11.30	41	-54	-22	6	5.49
	R	41	60	-22	6	4.15	41	48	-22	10	8.17	41	50	-20	10	4.14
	L											42	-56	-30	16	4.11
	R	42	60	-30	12	3.46										
Gyrus temporal inférieur /	L						37	-28	-50	-12	4.02					
Gyrus fusiforme	R						37	38	-66	2	5.33					
<u>Pariétales</u>																
Cortex somatosensoriel	L						3	-46	-10	30	10.26					
	R						3	56	-4	20	6.37					
Gyrus supramarginal / Sulcus pariétal inférieur	L	40	-48	-32	12	4.80	40	-48	-44	32	6.41					
	R						40	52	-28	16	8.77					
	R						40	38	-32	38	3.98					
Opercule rolandique	R						43	50	-6	14	6.52					
	L						43	-54	-12	14	7.67					
Précuneus	R						7	18	-80	34	4.12					
	R						7	20	-58	54	4.35					
Lobule pariétal supérieur	L						5	-14	-40	54	4.34					
	R						5	12	-28	44	4.22					
<u>Cortex insulaire</u>																
Insula	L	13	-44	0	0	5.21										
	R	13	44	-14	4	4.86	13	40	4	2	5.12					
Clastrum	R							26	2	22	4.66					
<u>Système limbique</u>																
Gyrus parahippocampique	R						34	22	2	-12	4.82					
Hippocampe	L							-32	-18	-20	6.82					
Cortex cingulaire antérieur ventral	L						24	-16	-10	48	4.57					
	R						24	8	14	30	5.03					
	R						24	12	4	50	4.67					
Cortex cingulaire antérieur dorsal	L						32	-8	16	32	3.73					
Cortex cingulaire retrosplénial	R						29	14	-50	8	5.67					
Cortex cingulaire postérieur dorsal	L						31	-14	-22	46	5.06					
	L						31	-4	-70	20	6.37					
<u>Thalamus</u>																
	L							-12	-20	8	3.51					
	R							24	-26	10	5.98					
Noyau latéral postérieur ventral	L							-14	-18	2	4.23					
Corps Génouillé médian	L							-14	-26	-6	3.21					
<u>Sous-corticales</u>																
Noyau caudé (tête)	L							-8	4	6	5.26					
	R							12	12	2	3.20					
Noyau caudé (corps)	R							18	10	22	3.40					
Putamen	L							-22	8	16	5.07					
	R							22	6	18	3.70					
Noyau rouge	L							-8	-24	-8	4.06					
Pallidus latéral	R							20	0	-8	5.17					
Pallidus médian	L							-18	-8	-2	3.83					
Substance noire	R							10	-24	-16	5.01					
<u>Cervelet</u>																
Culmen	R							8	-40	-6	4.25					
	L							-2	-38	-8	2.64					

Déclive	L	-6	-34	-26	5.48	
	L	-34	-76	-22	3.51	
	R	16	-66	-20	6.41	
<u>Occipitales</u>						
Cortex strié	L	17	-22	-62	6	5.93
	R	17	10	-72	10	5.34
Cortex extra-strié	L	18	-2	-74	-6	3.08
	L	19	-26	-58	-10	3.00

Table 2.4 : Coordonnées dans l'espace MNI des pics d'activation maximale observés pour les tâches de perception et de production des voyelles et communs à ces 2 tâches (analyse de conjonction). Résultats reportés à un seuil corrigé (FDR, $p < .05$) et une taille minimale des clusters de 25 voxels.

Effets principaux des degrés d'ouverture et d'arrondissement/antériorité

Trait d'ouverture: De manière générale, dans les deux tâches, l'effet principal d'ouverture a montré une croissance d'activation fonction du degré d'ouverture (voir Figure 2.10 et Table 2.5).

- Lors de la tâche de **perception**, cette tendance a été observée de manière bilatérale au niveau des gyri temporaux supérieur (incluant les parties postérieure et antérieure, BA 22, et le gyrus de Heschl, BA 41) et moyen (BA 21). Cet effet a été observé également au niveau des régions frontales inférieures postérieures (BA 44, 45) et dorsolatérales préfrontales (BA 9, 46) au niveau du gyrus angulaire bilatéral (BA 39) et de l'insula droite (BA 13).
- Lors de la tâche de **production**, cette croissance de l'amplitude du signal BOLD en fonction du degré d'ouverture a été observée au sein de l'aire de Broca (les parties operculaires et triangulaires dans les deux hémisphères), du cortex prémoteur gauche et du cortex dorsolatéral préfrontal bilatéral, ainsi qu'au niveau du précuneus (bilatéral), du gyrus supramarginal gauche et de l'aire visuelle extrastriée à gauche. En dehors de ces patterns d'activation, il est aussi à noter une déactivation de l'aire temporopolaire droite lors de la production des voyelles mi-ouvertes bien qu'une activation de cette région soit observée pour les autres types de voyelles.

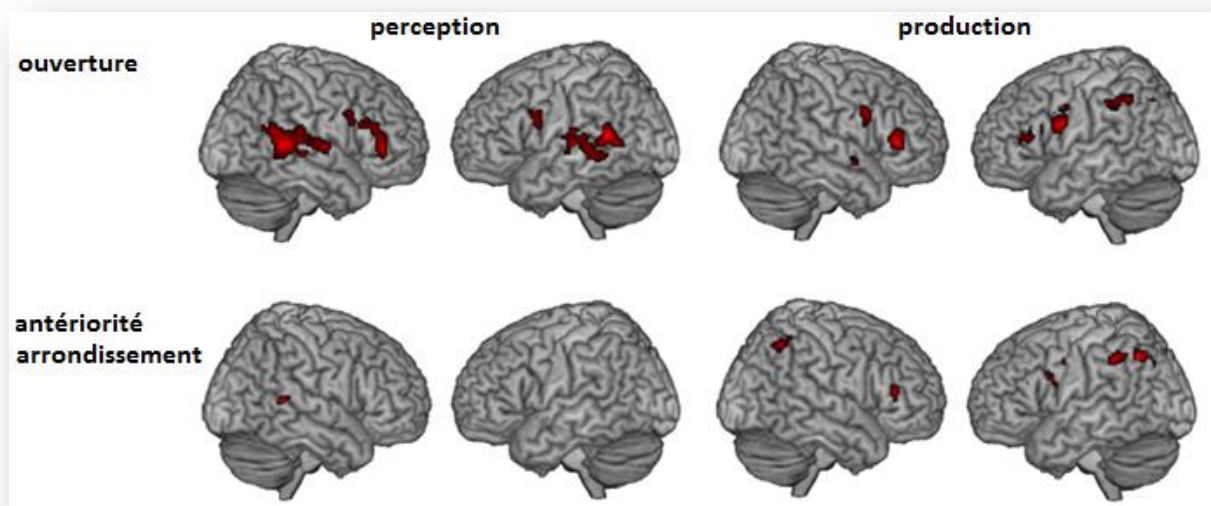


Figure 2.10 : Vue surfacique des régions cérébrales différenciellement activées lors des tâches de perception (gauche) et de production (droite) des voyelles en fonction des degrés d'ouverture (en haut) et d'arrondissement/antériorité (en bas) ($p < .001$ non corrigé et taille minimale des clusters de 25 voxels).

Traits d'arrondissement/antériorité : De manière générale, dans les deux tâches, nous observons une diminution d'activité en fonction de l'arrondissement des lèvres et de la position postérieure de la langue (voir Figure 2.10 et Table 2.5).

OUVERTURE			PERCEPTION							PRODUCTION							
Régions	H	BA	Coordonnées			F	Paramètres estimés			B A	Coordonnées			F	Paramètres estimés		
			MNI				F	Mi-F	Mi-O		MNI				F	Mi-F	Mi-O
			x	y	z						x	y	z				
<u>Frontales</u>																	
Gyrus frontal inférieur	L	44	-46	8	30	13.	-0.86	1.37	2.47	44	-48	4	24	20.6	-1.61	1.25	2.16
	R	44/	48	10	30	9.1	-0.41	-0.09	1.36	44	52	8	22	8.48	-0.57	1.30	0.89
	R	45	46	32	18	11.	-0.63	-0.10	1.30	45	-36	26	24	12.2	-0.94	-0.14	1.29
	L									45	54	34	6	20.8	-0.61	0.94	2.77
Cortex prémoteur	L									6	-42	0	28	19.0	-1.44	-0.02	2.13
Gyrus frontal moyen	L	9	40	12	24	12.	-0.65	-0.13	1.38	9	48	8	30	13.7	-0.85	0.28	1.90
Cortex dorsolatéral préfrontal	L	46	50	38	2	16.	-0.57	-0.21	1.73	46	-44	30	14	12.8	-1.04	0.52	1.70
<u>Temporales</u>																	
Aire temporopolaire	R									38	48	0	-10	9.14	0.59	0.84	-2.24
Gyrus temporal supérieur	L	22	-60	-48	12	21.	-0.24	-0.11	2.12								
- Partie postérieure du STG	R	22	54	-42	6	18.	-0.63	0.79	2.52								
- Partie antérieure du STG	R	22	52	-8	4	13.	1.14	0.37	2.52								
Gyrus temporal transverse	L	41	-44	-30	10	16.	0.45	0.67	2.66								
	R	41	44	-20	4	11.	0.99	0.38	2.06								
Gyrus temporal moyen	L	21	-54	-16	-2	11.	-0.07	-0.11	1.61								
	R	21	56	-24	-2	11.	-0.81	0.23	1.66								
<u>Pariétales</u>																	
Gyrus angulaire	L	39	-52	-56	6	13.	0.01	0.78	2.13								
	R	39	54	-48	16	13	-0.87	-0.13	1.46								
Lobule/sulcus pariétal inférieur	L									40	-42	-38	40	14.7	-1.06	0.35	1.59
Lobule pariétal supérieur	L									7	-12	-66	32	10.4	-0.65	1.27	1.40
	R									7	10	-74	30	10.4	-1.30	0.83	1.68
	L																
<u>Cortex insulaire</u>	L	13	44	-22	14	12.	0.85	-0.36	1.54	13	40	2	-4	10.1	0.08	1.47	-0.60
<u>Occipitales</u>																	
Cortex extrastrié	L									19	-30	-68	40	8.81	-1.76	-0.22	1.19

ARRONDISSEMENT/ANTERIORITE			PERCEPTION							PRODUCTION							
Régions	H	BA	Coordonnées MNI			F	Paramètres estimés			BA	Coordonnées MNI			F	Paramètres estimés		
			MNI				AN	AA	PA		MNI				AN	AA	PA
			x	y	z						x	y	z				
<u>Frontales</u>																	
Gyrus frontal inférieur	L									44	-46	10	28	10.40	1.26	-0.24	-0.36
	R									9/44	42	6	30	9.94	0.37	-0.48	-0.85
	R									45	54	34	6	8.77	1.37	1.27	0.46
	R									45/46	48	34	12	11.37	-0.68	0.63	1.95
Cortex prémoteur	L									6	-40	-2	38	8.75	0.99	0.24	-0.19
Gyrus frontal moyen	L									9	-36	24	28	16.40	0.64	-0.27	-0.63
<u>Temporales</u>																	
Gyrus temporal moyen	R	21	50	-42	8	9.41	1.45	0.83	0.26								
<u>Pariétales</u>																	
Gyrus/sulcus pariétal inférieur	L									40	-36	-60	42	18.20	0.59	-0.89	-1.33
	R									40	38	-56	46	12.96	0.38	-0.58	-1.17

Table 2.5 : Coordonnées dans l'espace MNI des régions différenciellement activées lors des tâches de perception et de production des voyelles en fonction des degrés d'ouverture et d'arrondissement/antériorité ($p < .001$ non corrigé et taille minimale des clusters de 25 voxels; F: voyelles fermées, mi-F: voyelles mi-fermées, mi-O: voyelles mi-ouvertes, AN: voyelles antérieures non arrondies, AA: voyelles antérieures arrondies, PA: voyelles postérieures arrondies).

- Lors de la tâche de **perception**, l'activation au niveau du gyrus temporal moyen droit est maximale pour les voyelles antérieures non arrondies et minimale pour les voyelles postérieures arrondies.
- Lors de la tâche de **production**, la même tendance est observée au niveau du gyrus frontal inférieur postérieur (BA 44 bilatéral et BA 45 à droite), au niveau du cortex prémoteur ventral gauche et de manière bilatérale au niveau du gyrus supramarginal (BA 40). On observe pour le seul cortex le cortex dorsolatéral préfrontal droit un effet inverse, avec une déactivation par

rapport à la ligne de base pour les voyelles antérieures non arrondies et une activation maximale pour les voyelles postérieures arrondies.

Effet de Répétition Suppression

L'analyse de groupe a montré plusieurs clusters sensibles à la modulation neuronale induite par la répétition des voyelles perçues et produites (voir Figure 2.11 et Table 2.6).

Perception

Les régions cérébrales sensibles à la répétition des voyelles perçues sont : l'aire motrice supplémentaire et le cortex cingulaire antérieure, le gyrus frontal moyen et le cortex dorsolatéral dans les deux hémisphères, l'aire de Broca dans sa partie operculaire bilatérale (ainsi que les parties triangulaire et orbitale à gauche), l'insula, le gyrus/sulcus temporal supérieur postérieur et moyen, le planum temporale gauche et le gyrus de Heschl droit. Parmi les structures pariétales, on observe le cortex pariétal supérieur dans sa partie antérieure (BA 5), le gyrus supramarginal et le sulcus intrapariétal, ainsi que le gyrus angulaire gauche. Enfin, un effet de RS est également observé dans le thalamus et le noyau rouge de l'hémisphère gauche.

Production

Les régions cérébrales montrant un effet de répétition suppression lors de la tâche de production sont : l'aire motrice supplémentaire et le gyrus frontal inférieur bilatéral (parties operculaire, BA 44, et triangulaire, BA 45), le cortex prémoteur ventral gauche, le cortex moyen et dorsolatéral préfrontal (BA 9, 10, 46). Parmi les aires temporales et pariétales, l'effet de RS est observé au sein du gyrus fusiforme (BA 37), des gyri temporaux inférieur gauche (BA 20) et moyen (BA 21), du gyrus/sulcus temporal supérieur droit (BA 22) ainsi que de manière bilatérale au niveau du lobule pariétal inférieur (sulcus pariétal inférieur, gyrus supramarginal et gyrus angulaire) et du lobule pariétal supérieur (précuneus). Parmi les structures sous-corticales, l'effet de RS concerne également le cortex cingulaire antérieur (ventral, BA 24, et dorsal, BA 34) et postérieur (ventral, BA 23, et dorsal, BA 31), l'insula droite, le putamen gauche, le cervelet (le culmen et le declive dans les deux hémisphères) et enfin le cortex extra-strié du lobe occipital.

Conjonction

Une diminution des activations au fur à mesure de la répétition des voyelles aussi bien perçues que produites a été observée au niveau du gyrus frontal inférieur gauche et du cortex prémoteur ventral adjacent (parties operculaire, BA 44, et triangulaire, BA 45), du cortex moyen et dorsolatéral préfrontal gauches (BA 9, 10, 46) et de l'aire motrice supplémentaire. Cet effet a également été observé au niveau du gyrus temporal supérieur/moyen (BA 22) et au niveau des régions pariétales gauche, au sein du lobule pariétal inférieur (gyrus angulaire et supramarginal) et supérieur (précuneus), tout comme au niveau du cortex cingulaire antérieur dorsal bilatéralement et du thalamus gauche.

D'autres régions supplémentaires montrent une sensibilité à la répétition de mouvements (gyrus frontal inférieur droit, parties operculaire, triangulaire et orbitale, et gauche dans la partie orbitale, ainsi qu'au niveau du cortex moyen et dorsolatéral préfrontal droite), sans avoir cependant survécu à un seuil corrigé au niveau du cluster.

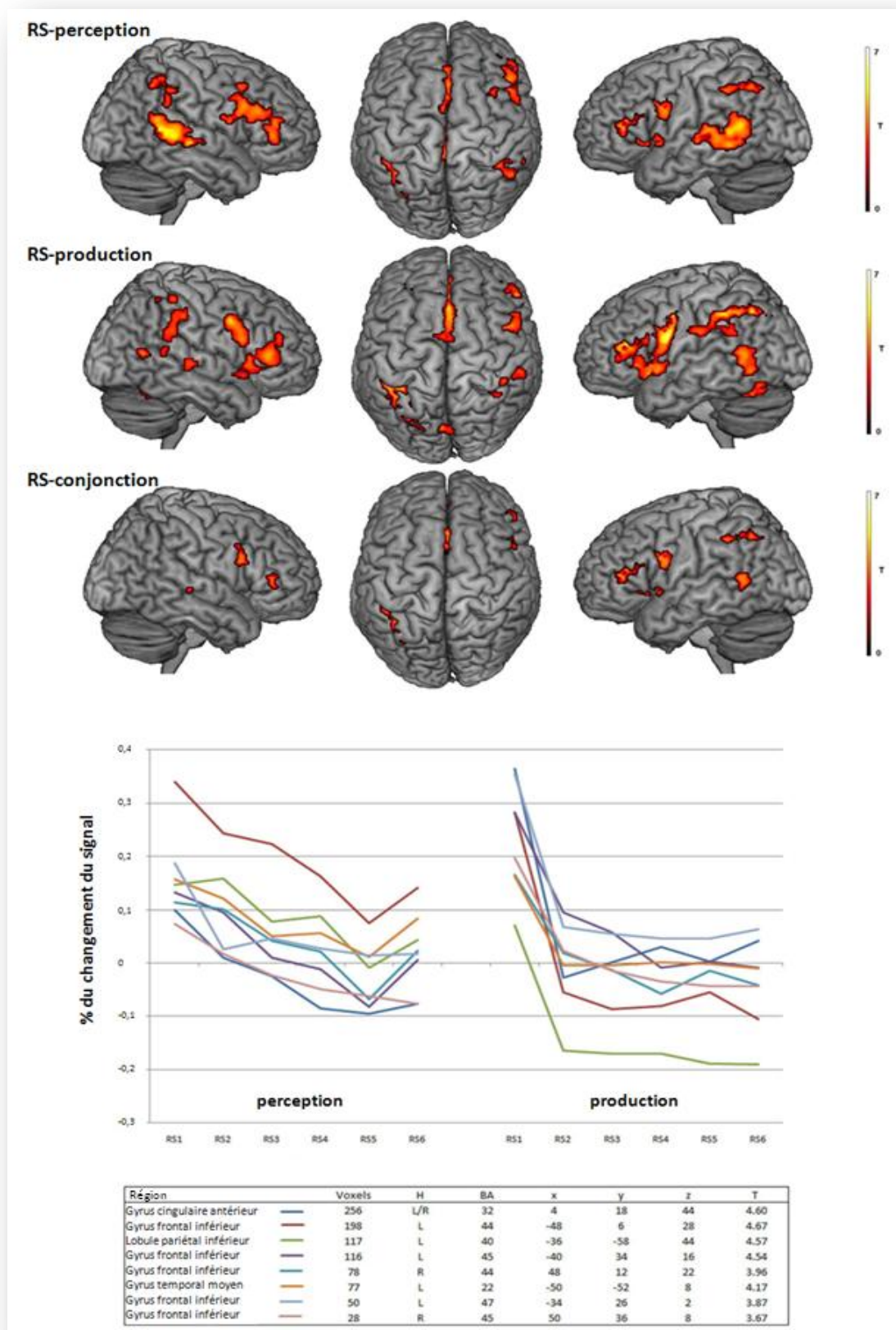


Figure 2.11 : Haut : Vue surfacique des régions cérébrales activées en fonction de l'ordre de répétition des voyelles perçues et produites ($p < .05$ FDR corrigé et taille minimale des clusters de 25 voxels). Bas : régions présentant un effet de RS lors de la perception et production des voyelles (décroissance de l'activité au cours des 6 répétitions; $p < .001$ non corrigé et taille minimale des clusters de 25 voxels).

Régions	H	BA	Perception				BA	Production				BA	Conjonction			
			x	y	z	T		x	y	z	T		x	y	z	T
Frontales																
Gyrus frontal inférieur	L	44	-38	10	28	4.53	44	-36	14	22	4.44	44	-48	6	28	4.67
	R	44	46	12	22	4.11	44	46	12	28	4.07	44	48	12	22	3.96
	L	47	-34	26	2	3.87	45	-40	34	16	5.63	47	-34	26	2	3.87
	L	45	-40	34	16	4.54	45	52	34	8	4.46	45	-40	34	16	4.54
	L						47	-34	28	2	5.41					
	R						47	46	20	-2	3.71	45	50	36	8	3.67
Cortex prémoteur	L						6	-42	-2	32	5.29					
Cortex préfrontal	R	9	34	10	40	3.55	9	44	6	36	5.20	9	46	14	26	3.90
	L											9	-38	6	26	3.81
Gyrus frontal moyen	L											10	-38	38	10	4.32
	R	10	36	38	14	4.36	10	28	36	18	3.46					
	R	8	34	14	44	3.66	10	36	42	24	3.32					
	L	8	-6	26	50	3.33										
Cortex dorsolateral préfrontal	R	46	46	16	28	4.34	9	54	10	36	3.45					
	L	46	-40	38	10	4.62						46	-40	36	12	4.41
	L	9	-46	8	32	5.02										
Gyrus frontal médial	R	6	6	28	38	4.24						9	2	44	20	3.20
Aire motrice supplémentaire	L	6	-4	2	60	3.63	6	4	16	54	4.28	6	0	6	56	3.23
Temporales																
Aire temporopolaire	L	22	-50	10	-2	3.66	38	50	16	-10	3.31	22	-50	10	-2	3.66
Gyrus temporal supérieur	L	22	-54	-52	12	5.16						22	-50	-52	8	4.17
	L	22	-56	-18	-2	3.71										
	R	22	50	-40	6	6.22	22	58	-24	0	3.55	22	58	-24	0	3.55
	R	22	56	-16	0	3.26	22	52	-44	12	3.01					
Gyrus temporal transverse	R	41	60	-26	8	3.24										
Gyrus temporal inférieur	R						20	34	-44	-20	3.70					
Gyrus fusiforme	L						37	-36	-62	2	3.10					
	R						37	40	-50	-20	3.86					
Gyrus temporal moyen	L	21	-60	-42	-4	4.58	21	-56	-60	2	3.31					
	R	21	56	-24	-2	5.21										
Pariétales																
Cortex somatosensoriel	L						3	-40	-16	34	3.95					
	R						2	34	-40	46	4.00					
	R						2	42	-38	54	3.47					
Gyrus angulaire	L	39	-48	-50	12	5.13	39	-48	-54	6	4.61	39	-32	-68	42	3.27
	R						39	46	-60	10	4.00					
Gyrus supramarginal / Sulcus pariétal inférieur	L	40	46	-40	42	3.56						40	-36	-58	44	4.57
	L	40	-32	-68	42	3.27						40	-38	-50	42	3.84
	R	40	54	-50	18	4.62										
Gyrus supramarginal Sulcus pariétal inférieur	L	40	-44	-42	34	3.29										
	L	40	-36	-56	44	4.96	40	-44	-38	40	5.57					
	L	40	-46	-40	42	3.80	40	36	-44	38	3.76					
	R	40	40	-44	46	3.74	40	60	-38	26	3.45					
Opercule rolandique																
Précuneus	L						7	-36	-46	48	5.40	7	-10	-68	38	3.23
	L						7	-6	-74	36	4.47					
	L						7	-22	-62	32	4.22					
	R						7	10	-72	28	4.09					
	R						7	36	-48	50	2.78					
Lobule pariétal supérieur	L	5	-6	-42	50	3.00										
	R	5	8	-42	50	3.84										
Cortex insulaire																
Insula	L	13	-40	-24	16	3.04										
	R	13	34	28	18	4.67	13	34	20	6	4.24					
Clastrum	R							30	12	6	3.57					
Système limbique																
Cortex cingulaire antérieur ventral	L	24	-4	-6	34	3.27	24	-2	14	26	5.22					
	R	24	2	-16	38	3.99	24	4	28	24	5.10					
Cortex cingulaire antérieur dorsal	L	32	-6	24	40	4.59	32	-6	18	28	4.81	32	-2	22	40	4.15
	R	32	4	18	44	4.69	32	4	38	16	4.61	32	4	18	44	4.60
Cortex cingulaire retrosplenial	R	29	-10	-44	8	3.94										

		30	0	-48	20	3.76											
Cortex cingulaire postérieur dorsal	L	31	-8	-44	30	3.36	31	-14	-64	30	4.27						
Cortex cingulaire postérieur ventral	L	23	-2	-34	30	4.56	23	-6	-24	30	3.76	23	-2	-32	26	3.60	
	L						23	-4	-40	22	3.08						
	R						23	8	-26	34	3.10						
Thalamus	L	-6	-14	-2	3.56								-6	-10	8	3.30	
<u>Ganglions de la base et tronc cérébral</u>																	
Noyau caudé (corps)	R						18	16	8	4.39							
Putamen	L						-18	4	6	5.09							
Noyau rouge	L	-4	-14	-6	3.56												
<u>Cervelet</u>																	
Culmen	L						-28	-56	-30	3.65							
	L						-14	-64	-16	3.23							
	R						2	-60	-26	3.43							
	R						12	-62	-16	3.40							
Declive	L						-40	-58	-24	4.20							
	R						38	-58	-22	3.87							
<u>Occipital es</u>																	
Cortex extrastré	L						19	-46	-58	-6	4.40						

Table 2.6 : Coordonnées dans l'espace MNI des régions présentant un effet de RS lors de la perception et production des voyelles (décroissance de l'activité au cours des 6 répétitions; $p < .05$ FDR corrigé et taille minimale des clusters de 25 voxels) et régions communes aux deux tâches ($p < .001$ non corrigé et taille minimale des clusters de 25 voxels).

2.2.4 Discussion

La présente étude IRMf avait pour but de déterminer les régions neurales sensibles à un effet de répétition suppression lors de l'écoute ou de la production répétée d'une même voyelle. En accord avec nos hypothèses, un effet d'adaptation phonémo-spécifique a été observé au sein des régions de la voie dorsale de l'hémisphère gauche, incluant le gyrus frontal inférieur et le cortex prémoteur ventral adjacent, les régions postérieures du gyrus/sulcus temporal supérieur et du gyrus temporal moyen et, enfin, le lobule pariétal inférieur et le sillon intrapariétal.

Ces résultats suggèrent des mécanismes sensorimoteurs adaptatifs communs lors des traitements auditifs et des processus de contrôle moteur de la parole liés au but des actions vocaliques perçues et produites et démontrent de fait l'existence d'un couplage fonctionnel entre systèmes de perception et de production de la parole.

Réseaux des systèmes de perception et de production

Les activations neuronales liées à la perception et production des 9 voyelles et traits phonétiques correspondants (non reliées à l'effet de RS) sont similaires à celles observées lors de l'étude présentée dans le Chapitre 1.2. Bien évidemment, le degré moindre d'activation observé pour certaines régions (par exemple au niveau du gyrus frontal inférieur et du cortex prémoteur ventral adjacent) s'explique par le protocole d'adaptation et l'effet de répétition suppression. L'objectif de cette étude n'était pas de mettre en valeur de manière optimal le réseau neural des systèmes de perception et production des voyelles, comme lors de la précédente étude, mais de déterminer les régions impliquées dans les processus sensorimoteurs adaptatifs phonémo-spécifiques en dehors de processus moteurs, sensoriels ou exécutifs plus généraux.

Les analyses des activations moyennes des voyelles perçues révèlent des activations bilatérales temporales, correspondant aux aires auditives primaires, secondaires et associatives, ainsi que l'activation du cortex insulaire. On observe de plus dans l'hémisphère gauche, une activation de la partie operculaire du gyrus frontal inférieur ainsi que du gyrus supramarginal. Ainsi, malgré un degré d'activité moindre dû à l'effet de RS, nous observons comme dans l'étude précédente l'implication

de la voie dorsale lors de la perception des voyelles. En production, de larges activations bilatérales, corticales, sous-corticales et cérébelleuses démontrent l'implication de régions dédiées au contrôle moteur orofacial en général (voir Chapitres 1.1 et 2.1) et à la production de voyelles en particulier (voir Chapitre 1.2). Enfin, l'analyse de conjonction a permis de mettre en évidence des régions temporales (incluant les cortex primaire, secondaire et associatif) communes aux deux tâches. Le fait que les régions frontale et pariétale n'apparaissent pas lors de l'analyse de conjonction est plus que probablement lié à l'effet adaptatif observé dans ces régions alors que les régions temporales n'ont pas toutes subi un tel effet de RS (voir ci-dessous).

Représentations différenciées entre voyelles

Tout comme lors de la précédente étude (voir Chapitre 1.2) et en accord avec l'étude récente d'Obleser et al. (2010), nous ne trouvons pas de différences topographiques entre voyelles. En revanche, une modulation d'activité de certaines régions est observée en fonction des traits phonétiques perçus et produits. Notamment, comme lors de l'étude précédente, une implication différenciée de certaines régions frontales et pariétales motrices est de nouveau reportée pour le trait d'ouverture lors de la tâche de production. De manière intéressante, cet effet est également observé lors de la perception des voyelles pour les régions frontales inférieures et temporales auditives.

Trait d'ouverture

En accord avec les résultats de l'étude précédente, nous observons une augmentation d'activité de certaines régions cérébrales lors de la production et de la perception des voyelles mi-ouvertes par rapport aux voyelles fermées et mi-fermées.

Production. Cette augmentation d'activité pour les voyelles mi-ouvertes est observée en production au niveau du gyrus frontal inférieur et du cortex dorsolatéral préfrontal de manière bilatérale et, dans l'hémisphère gauche, du cortex prémoteur, des lobules pariétaux inférieur et supérieur et du cortex strié. Comme supposé lors de la précédente étude (voir la discussion du Chapitre 1.2), les activations prémotrices accrues pourraient s'expliquer par la moindre naturalité de production des voyelles mi-ouvertes en contexte isolé et, de là, des traitements prémoteurs plus importants. De même, l'activité supérieure observée pour ces voyelles au niveau des régions pariétales peut être associée à une implication accrue de processus de vérification en ligne des conséquences somatosensorielles prédites et réelles (Guenther, 2006; Guenther et Vladusich, 2012). Dans cette ligne de réflexion, l'implication accrue du cortex dorsolatéral préfrontal bilatéral n'apparaît pas étonnante puisqu'elle pourrait refléter une plus grande difficulté de réalisation pour ces voyelles⁵². L'implication du cortex strié pourrait enfin s'expliquer par l'utilisation de consignes visuelles sans doute moins connues des participants (/ɛ/, /œ/ et /ɔ/). De manière intéressante, il est enfin à noter que nous n'observons pas de modulations similaires d'activité au niveau du cortex auditif, ce qui appuie indirectement l'existence de boucles correctives distinctes des retours somatosensoriel et auditif.

Perception. De manière tout à fait intéressante, une telle augmentation d'activité pour les voyelles mi-ouvertes est également observée lors de la tâche de perception. Des activations accrues de larges parties des régions auditives temporales supérieures et moyennes, s'étendant dans l'hémisphère droit au cortex insulaire postérieur, témoignent de traitements auditifs plus importants lors de la perception et de l'interprétation de ces voyelles. Le gyrus angulaire (BA 39) en tant que sous-région du lobule pariétal inférieur, serait impliqué au même titre que le gyrus supramarginal dans la perception des sons de parole (voir par exemple Turkeltaub et Coslett, 2010). De manière plus

⁵² Bien qu'activée lors de la production des autres types de voyelles, une déactivation de l'aire temporopolaire droite lors de la réalisation des voyelles mi-ouvertes est également observée. Nous n'avons cependant pas d'hypothèses claires quant à ce résultat.

spécifique, pour certains auteurs, le gyrus angulaire aurait plutôt un rôle d'extraction sémantique (Obleser et al., 2007; pour une revue voir aussi Price, 2010). Pour d'autres auteurs (par exemple, Booth et al., 2007), le gyrus angulaire serait plutôt impliqué lors du décodage visuel des lettres, en reliant les étapes de traitement visuel du cortex strié avec les informations phonologiques qui sont proposées être codées au niveau du gyrus supramarginal. A partir de là, une interprétation pourrait être celle d'une visualisation irrépressible des symboles visuels pour les voyelles mi-ouvertes, symboles appris par les participants seulement très récemment avant d'effectuer l'expérience. Il est donc possible que cette association soit particulièrement évoquée pour des apprentissages récents, maîtrisés, mais non automatisés. Comme lors de la tâche de production, l'activation du cortex dorsolatéral préfrontal pour les voyelles mi-ouvertes peut être interprétée comme une focalisation de ressources attentionnelles envers des sons de parole non typiquement perçus dans un contexte isolé, donc plus difficilement à reconnaître. Cette interprétation est d'ailleurs renforcée par l'observation d'activation accrue du gyrus frontal inférieur, suggérant la mise en œuvre de mécanismes de simulation motrice plus importants dans le cas d'une perception moins naturelle (Schwartz et al., 2012).

Traits d'arrondissement et d'antériorité

Contrairement à l'étude présentée dans le Chapitre 1.2, nous observons ici des activations différenciées de certaines régions cérébrales en fonction des traits d'arrondissement et d'antériorité des voyelles produites et perçues, consistant en une diminution d'activité en fonction de l'arrondissement des lèvres et de la position postérieure de la langue. Néanmoins, nous n'avons pas ici d'explications claires quant à ces différences d'activités.

Production. On observe une diminution d'activité du gyrus frontal inférieur postérieur bilatéral, du cortex prémoteur gauche et du gyrus supramarginal entre voyelles antérieures non-arrondies, antérieures arrondies et postérieures arrondies.

Perception. La seule observation significative consiste en une baisse d'activation fonction de l'arrondissement des lèvres et de la position postérieure de la langue au niveau du gyrus temporal moyen droit.

Répétition Suppression

L'objectif principal de la présente étude était de déterminer les populations neuronales cérébrales spécifiquement sensibles au codage phonémique aussi bien en perception qu'en production et ce, en utilisant un effet d'adaptation phonémo-spécifique consistant en une écoute ou une réalisation répétée six fois d'une même voyelle toutes les 10 secondes⁵³.

Production

Codage du but vocalique/phonémique au sein de la voie dorsale.

Un effet de RS a été observé lors de la production répétée des voyelles au niveau du gyrus frontal inférieur et du cortex prémoteur ventral adjacent, du lobule pariétal inférieur et du sillon intrapariétal et, enfin, des parties postérieures du gyrus/sulcul temporal supérieur et du gyrus temporal moyen. L'effet d'adaptation observé au sein de ces régions de la voie dorsale suggère ainsi un réseau fronto-pariéto-temporal de codage du but des actions vocaliques/phonémiques, réseau distribué entre régions associatives codant les buts articulatoires, les buts auditifs et

⁵³ Il est à noter que notre paradigme a permis d'observer un effet de RS malgré cet intervalle de présentation de 10sec. De précédentes études utilisant cette méthode d'adaptation ont montré un tel effet de RS avec un intervalle inter-stimulus plus rapide : 50ms (Joanisse et al., 2007), 150 ms (Vaden, Muftuler et Hickok, 2010), 300ms (Peeva et al., 2010), 500ms (Raizada et Poldrack, 2007), 3000ms (Hasson et al., 2007). Un intérêt méthodologique de notre étude est donc de démontrer un effet de RS avec une distance temporelle de répétition de 10 secondes (voir également Chapitre 2.1).

somatosensoriels et les représentations multimodales du corps dans l'espace.

- Cortex prémoteur ventral et gyrus frontal inférieur (incluant les parties operculaire, triangulaire et orbitale) : Nous l'avons vu, l'aire de Broca est traditionnellement impliquée dans la production de la parole et du langage, et notamment pour la partie operculaire et le cortex prémoteur ventral adjacent dans le codage des paramètres et représentations motrices sous-jacents (ces deux régions ont été proposées contenir une "carte des sons de parole"; Guenther, Ghosh et Tourville, 2006; Guenther et Vladusich, 2012; voir Chapitre 1.2 pour plus de détails). L'idée que le codage des paramètres articulatoires des unités de parole au niveau des régions frontales inférieures fasse partie du codage plus général du but de l'action n'est pas nouvelle. Comme mentionné précédemment, des enregistrements neurophysiologiques unicellulaires effectués chez le primate non humain (Rizzolatti et al., 1988; Fogassi et al., 2005; Gardner et al., 2007; Rozzi et al., 2008; Umiltà et al., 2008; Bonini et al., 2010, 2011), chez des patients souffrant d'apraxie idéomotrice (Haaland, Harrington et Knight, 2000) et des études IRMf utilisant un protocole d'adaptation chez l'homme (Dinstein et al., 2007; Chong et al., 2008; Kilner et al., 2009; Hamilton et Grafton, 2009; Peeva et al., 2010) ont fourni de nombreuses évidences en faveur d'un codage du but de l'action dans ces régions et, notamment, le codage des paramètres articulatoires et du plan moteur de l'action exécutée.
- Lobule pariétal inférieur et sillon intra-pariétal : L'idée que les régions pariétales inférieures et le sulcus intra-pariétal soient impliquées dans la localisation des buts d'action dans l'espace n'est pas nouvelle, notamment l'idée que ces régions coderaient les connaissances pragmatiques liées aux actions (voir par exemple, Goodale et Milner, 1992; Gallese et al., 1994; Sakata et al., 1995; Jeannerod, 1997). En accord avec ces études, Haaland, Harrington et Knight (2000) ont montré que les patients souffrant d'apraxie idéomotrice⁵⁴ auraient en effet souvent des lésions au niveau des régions entourant le sulcus intrapariétal (BA 7, 39, 40). En plus du rôle de ces régions dans les mécanismes d'intégration motori-somatosensoriels tel que suggéré dans l'étude présentée dans le Chapitre 2.1, il a été récemment montré que ces régions sont impliquées dans l'intention d'agir, par exemple lors d'une prise de parole (Carota et al., 2010) où un rôle de surveillance multimodale de l'action et de l'intention d'agir est attribué à ces régions pariétales inférieures. Sur ce point, d'autres études ont montré que ces régions seraient impliquées de manière plus générale dans le choix du but de l'action (Hamilton et Grafton, 2006) et qu'une stimulation de ces régions peut induire des envies, voire des illusions de mouvements de parties du corps selon une certaine somatotopie grossière (Desmurget et al., 2009; Desmurget et Sirigu, 2009). De là, il serait possible de conclure de nos résultats que le lobule pariétal inférieur et le sillon intra-pariétal seraient impliqués dans les mécanismes d'intégration sensorimotrice et de choix de l'action vocale⁵⁵.
- Gyrus/sulcus temporal supérieur (BA 22) et moyen (BA 21) postérieurs : En accord avec l'implication de ces régions dans l'encodage des représentations phonémiques (voir Chapitre 1.2 pour plus de détails), une étude récente a montré un effet d'adaptation neuronale lors de la production répétée de mêmes phonèmes au niveau du gyrus temporal supérieur postérieur (Peeva et al., 2010), bien qu'aucun effet de RS n'ait été trouvé au niveau des aires pariétales ou frontales inférieures.

⁵⁴ Contrairement à l'apraxie idéatoire (considérée comme une désorganisation de la représentation mentale des gestes à exécuter, soit une difficulté à réaliser un mouvement composé d'un ensemble de sous-mouvements élémentaires organisés) qui serait liée à une lésion bilatérale occipito-pariétale, l'apraxie idéomotrice est un signe clinique qui décrit une incapacité à effectuer un mouvement sur consigne ou de mimer une action (Liepmann, 1908). Elle est souvent mise en parallèle avec des lésions au niveau des régions frontales inférieures et pariétales postérieures, le plus souvent dans l'hémisphère gauche.

⁵⁵ En accord avec cette hypothèse, une étude TMS récente (Shum et al., 2012) a montré que la stimulation de la partie dorsale du gyrus supramarginal diminuait les mécanismes de compensation articulatoire lors d'une tâche de production de parole avec modification en ligne du feedback acoustique par rapport à une condition contrôle.

Autres structures sensibles à l'effet de RS

D'autres structures cérébrales ont été également sensibles à la production répétée des voyelles : Le cortex dorsolatéral préfrontal et moyen, l'aire motrice supplémentaire, le cortex cingulaire antérieur, les cortex somatosensoriel primaire et le précuneus, ainsi que l'insula droite, le putamen, le cervelet (culmen et declive dans les deux hémisphères) et les aires visuelles striées et extra-striées.

Le cortex dorsolatéral préfrontal est une aire reliée aux fonctions exécutives et à l'intégration des informations sensorielles et mnésiques, en tant que telle, elle est à la fois impliquée dans la mémoire de travail et dans la régulation des plans moteurs. L'aire motrice supplémentaire et la partie antérieure du cortex cingulaire sont généralement considérées comme des régions permettant le contrôle volontaire du déclenchement ou la suppression des actes vocaux (Jürgens, 2002, 2009), tandis que le cortex somatosensoriel primaire traite les informations proprioceptives des articulateurs supralaryngés et laryngés et des poumons. Toutes ces structures exécutrices montrant un effet de répétition suppression témoignent d'une implication similairement dans chaque occurrence des actes répétés de production de voyelles. Quant au précuneus, il a été montré que chez l'humain, le lobule pariétal supérieur postérieur serait impliqué dans des interactions visuo-motrices, en terme d'imagerie visuelle de l'action (Lamm et al., 2007) ou de codage proprioceptif de l'action (Lestou, Pollick et Kourtzi, 2008).

Pour certains auteurs (par exemple, Wise et al., 1999; Dronkers et al., 1996, 2007), le rôle de l'insula antérieure dans la production de la parole consiste en la programmation articulatoire des gestes de parole, bien que d'autres auteurs attribuent ce rôle explicitement au gyrus frontal inférieur (par exemple, Hillis et al., 2004; Bonilha et al., 2006) et suggèrent que l'apraxie de la parole serait due aux disconnexions de la matière blanche entre régions temporales supérieures et de l'aire de Broca mais non à cause des lésions de la matière grise insulaire (Bonilha et al., 2009).

Le cervelet étant le principal centre de régulation de la fonction motrice, une de ses fonctions consiste à assurer la coordination motrice en effectuant une organisation temporo-spatiale à travers les boucles cortico-cérébelleuses et en jouant le rôle d'interface avec les régions sensorimotrices corticales. Pour cette raison, certains auteurs considèrent le cervelet comme le siège des "modèles internes" des actes moteurs (par exemple, Kawato, 1999), incluant également des actions orofaciales et liées à la parole (Guenther, Hampson et Johnson, 1998; Bohland et Guenther, 2006; Guenther, 2006; Guenther, Ghosh & Tourville, 2006).

Le putamen gauche et le noyau caudé droit sont également sensibles à l'effet de RS. Il est à noter que ces structures sont impliquées dans les boucles cortico-basales qui permettent la modulation de l'articulation et de la phonation (par exemple, Fabbro et al., 1996).

Enfin, de manière non surprenante, les aires visuelles (cortex extra-strié, gyrus fusiforme et temporal inférieur) ont subi un effet d'adaptation, sans doute dû à la vision répétitive et la lecture et compréhension des mêmes consignes visuelles des voyelles à produire.

Perception

Codage du but vocalique/phonémique au sein de la voie dorsale.

Les régions cérébrales dont l'activité BOLD décroît au fur et à mesure de l'écoute d'un même type de voyelle sont principalement les régions auditives temporales postérieures et le gyrus frontal inférieur.

- Gyrus/sulcus postérieur moyen, planum temporale gauche, gyrus de Heschl droit : En regard des désaccords concernant l'emplacement des populations neuronales responsables de l'encodage auditif phonologique (Belin et Zatorre, 2000; Scott et Johnsrude, 2003; Obleser et al., 2006; Hickok et Poeppel, 2007; Rauschecker et Scott, 2009; Rauschecker, 2011; voir aussi la discussion de l'étude 1.2.), nos résultats montrent un effet d'adaptation phonémique au niveau des régions auditives postérieures. Ce résultat corrobore de plus les quelques études IRMf récentes utilisant le même paradigme de répétition suppression et montrant la sensibilité de ces mêmes régions

au codage phonologique, lors d'une tâche de perception catégorielle (Joanisse, Zevin et McCandliss, 2007), de perception audio-visuelle syllabique (Hasson et al., 2007) ou lors d'une tâche de détection de phonèmes au sein d'un mot (Vaden, Muftuler et Hickok, 2010).

- En ce qui concerne la partie operculaire du gyrus frontal inférieur, qui correspond aux représentations articulatoires de l'action, nos résultats apparaissent cohérents avec deux études qui ont effectivement montré par des effets d'adaptation neuronale que l'encodage des percepts syllabiques nécessitant un traitement phonologique (percepts fusionnels audiovisuels, Hasson et al., 2007; perception catégorielle, Raizada et Poldrack, 2007) se fait également au niveau de l'aire de Broca.
- L'effet de RS est également observé au niveau du lobule pariétal inférieur, que l'on sait être impliqué depuis les travaux d'Alan Baddeley (1986) et de Paulesu, Frith et Frackowiak (1993) dans les processus de stockage phonologique en mémoire de travail verbale. Une hypothèse est donc que cette région code les buts phonémiques multimodaux. En accord avec cette proposition, Raizada et Poldrack (2007) montrent une amplification sélective des réponses neuronales au niveau du gyrus supramarginal en réponse à deux stimuli qui dépassent la frontière phonétique mais pas si la même différence acoustique a lieu au sein de la même catégorie phonologique. Aussi, une récente étude TMS (Hartwigsen et al., 2010) montre que la perturbation de cette région induit une baisse de performance dans des tâches de décisions phonologiques lexicales. Rappelons enfin le rôle de la partie dorsale du lobule pariétal inférieur et du sillon intrapariétal dans les mécanismes d'intégration motori-somatosensoriels tel que suggéré dans l'étude présentée dans le Chapitre 2.1.

Pris ensemble, ces résultats suggèrent ainsi un codage distribué des représentations et des buts des actions vocaliques/phonémiques impliquant un codage des représentations articulatoires dans la partie operculaire du gyrus frontal inférieur, auditives dans la partie postérieure du gyrus/sulcus temporal supérieur et somatosensoriels dans le sillon intrapariétal ainsi qu'un possible stockage multimodal des représentations vocaliques/phonémiques dans le gyrus supramarginal.

Autres structures sensibles à l'effet de RS.

D'autres structures cérébrales montrent également un effet d'adaptation lors de la perception répétée des voyelles : le cortex préfrontal dorsolatéral et moyen, l'insula, le cortex cingulaire et l'aire motrice supplémentaire, le cortex pariétal antérieur, le thalamus et le noyau rouge. Le rôle possible de la plupart de ces régions a été précédemment discuté. Par rapport à la décroissance d'activité du cortex dorsolatéral préfrontal (une structure impliquée dans la mémoire de travail verbale et les fonctions exécutives), celle-ci pourrait découler d'une décroissance de l'attention portée à l'écoute des voyelles présentées de manière répétitive. Enfin, il n'est pas étonnant de trouver des effets d'adaptation au niveau du thalamus puisque cette structure sous-corticale est composée, entre autre, de nombreux noyaux relais, dont notamment le corps génouillé médian organisé de manière tonotopique qui reçoit des afférences du colliculus inférieur provenant des nerfs cochléaires, et projette les informations acoustiques sur l'aire auditive primaire au niveau du gyrus temporal supérieur.

Conjonction

L'analyse de conjonction a permis de mettre en évidence les régions communes sensibles à l'effet de RS lors des tâches de perception et de production des voyelles : le gyrus frontal inférieur dans les parties operculaire, triangulaire et, à gauche, orbitale, les gyri temporal supérieur/moyen postérieurs et le lobule pariétal inférieur.

En somme, nous observons que les représentations neurales des voyelles (qu'elles soient perçues ou produites) semblent être distribuées à travers des régions temporales (codant les représentations ou buts auditifs), frontales inférieures postérieures (codant les représentations ou buts articulatoires) et pariétales inférieures (codant les buts vocaliques multimodaux). Ces résultats suggèrent,

qu'indépendamment de la tâche, les voyelles seraient représentées dans le cerveau de manière similaire en termes de codes auditifs, articulatoires et multimodaux. Ce résultat va dans le sens d'un système de codage commun pour les buts vocaliques perçus et produits.

Parmi les autres structures montrant des effets de répétition suppression pour dans deux tâches, les autres régions s'adaptant à la fois pour les voyelles perçues et produites incluent le cortex dorsolatéral préfrontal, l'aire motrice supplémentaire et le gyrus cingulaire antérieur dorsal, tout comme le thalamus gauche et le lobule pariétal supérieur (pour une discussion de ces fonctions, voir ci-dessus).

Patterns de décroissance.

Bien que ces régions montrent un effet de RS aussi bien lors de la perception que de la production répétée des voyelles, leurs patterns de décroissance diffèrent cependant entre les deux tâches mais ce de manière similaire pour toutes les régions, ce qui suggère fortement un effet de la tâche. Effectivement, si dans la tâche de production le signal BOLD chute brusquement à partir de la deuxième occurrence puis reste à un niveau stable par la suite, on observe lors de la tâche de perception une décroissance progressive de la première à la cinquième occurrence, et un rebond à la dernière répétition.

Il est important de rappeler que les participants savaient que les stimuli se répétaient six fois et pouvaient ainsi prédire les cinq occurrences suivant la première voyelle nouvellement présentée. Lors de la tâche de production, la chute brusque initiale pourrait être liée à l'élaboration d'un programme moteur dès l'apparition de la consigne lors de la première occurrence, qui n'a pas besoin d'être recréé six fois. Lors de la tâche de perception, il est possible qu'un facteur supplémentaire attentionnel se mêle à l'effet de RS⁵⁶. Ainsi, ces différences d'expectation et d'attention pourraient expliquer les patterns de décroissance observés lors des deux tâches.

En appui de cette hypothèse, Summerfield et al. (2008) ont suggéré que l'effet de RS serait une conséquence des attentes perceptives du sujet. Ils ont montré que la perception répétée d'un stimulus visuel réduit la réponse BOLD plus fortement lorsque cette répétition était prévisible que lorsqu'elle ne l'était pas (l'effet de RS était alors atténué). Cela a également été montré dans la modalité auditive. Todorovic et al. (2011) ont récemment effectué une étude MEG, dans laquelle ils ont montré qu'un effet d'expectation top-down augmente l'effet de RS au niveau du cortex auditif pour la présentation répétée des sons synthétisés. Quand les participants savaient qu'il y aurait répétition des tons perçus, l'effet RS était plus robuste que lorsque cette répétition intervenait de manière imprévue. En revanche, Rahnev, Lau et de Lange (2011) ont reporté une conclusion différente quant à l'effet d'expectation sur l'adaptation neuronale du cortex dorsolatéral préfrontal dans une tâche visuo-motrice. En effet, la survenue attendue d'un stimulus visuel améliorait d'une part la performance tout en augmentant l'activation du cortex dorsolatéral préfrontal. Cependant, les auteurs suggèrent que cette réponse amplifiée au niveau du cortex dorsolatéral préfrontal soit liée à l'intégration de la prévision des stimuli visuels aux processus de décision perceptive et de réponse manuelle.

Lors de la tâche de perception, le rebond de l'activité BOLD lors de l'occurrence finale pourrait témoigner d'un regain attentionnel, possiblement lié à l'attente d'un nouveau stimulus. Par contraste, lors de la tâche de production, il est important de garder l'attention la première fois pour apprendre l'instruction de la voyelle à produire les six prochains essais et, par la suite, la cible visuelle sert uniquement d'indicateur de l'onset de la production. Ce résultat suggère qu'un facteur supplémentaire, l'attention, module l'effet de RS. Cependant, les auteurs d'une récente étude (Kok et al., 2012) proposent une séparation des mécanismes prédictifs des mécanismes attentionnels. Dans leur étude, ils croisent ces deux facteurs, et montrent que l'attention inverse l'effet de RS. Selon leurs résultats, l'effet RS est fortement influencé par la prédiction d'une occurrence, plus qu'en situation

⁵⁶ Rappelons, que la tâche de perception était passive et durait 30 minutes. De plus, les voyelles étant répétées par trains de six trials, dans une minute les participants percevaient six fois le même type de voyelles, une toutes les dix secondes.

où l'on ne peut prédire la survenue d'un stimulus. L'effet RS serait donc maximal lorsqu'une répétition est attendue. Les travaux de Valentini (2011) appuient cette possibilité et suggèrent que l'effet RS reflète une combinaison de traitements hiérarchiques bottom-up et d'attentes perceptives top-down⁵⁷.

Limites de l'étude

Les limites inhérentes à la présente étude sont pour la plupart similaires à l'étude présentée dans le Chapitre 2.1. Ainsi, les paramètres musculaires et cinématiques des mouvements orofaciaux n'ont pas été contrôlés lors de la production des voyelles. Il est cependant important de noter qu'une variation des paramètres de force musculaire et cinématiques est fortement probable lors de la répétition de voyelles et, plus important encore, que l'écoute répétée de chaque voyelle consistait en une présentation de six occurrences distinctes d'une même voyelle. De fait, les effets de RS observés commun aux deux tâches suggèrent qu'ils ne dépendent pas de l'exécution ou de l'écoute précise et identique d'une voyelle d'un essai à l'autre mais bien du codage du but vocalique.

Egalement, dans le cadre d'un possible couplage fonctionnel entre systèmes de perception et de production, une des critiques du groupe de Caramazza (Lingnau, Gesierich et Caramazza, 2009) face aux effets de RS précédemment observés pour des actions manuelles perçues ou produites concerne la nécessité de démontrer un effet de RS cross-modal. Par rapport à l'étude présente, cela consisterait à démontrer un effet de RS aussi bien lors de l'écoute ou de la production répétée de voyelles mais également lors de la production puis de la perception d'une même voyelle et inversement. Si cette critique nous apparaît tout à fait pertinente (voir d'ailleurs Kilner et al., 2009 pour une démonstration de cette effet de RS cross-modal lors d'actions manuelles perçues et produites), le fait de mêler ainsi des tâches motrice et perceptive pourrait entraîner selon nous la possibilité non négligeable de mécanismes de simulation motrice et de mémoire de travail accrus, voire conscients, lors de la perception des voyelles. Ainsi, en vue de minimiser un tel biais possible, la tâche de perception auditive a toujours précédé la tâche de production dans la présente étude.

2.2.5 Conclusion

Cette étude constitue une première exploration des corrélats neuronaux des mécanismes sensorimoteurs adaptatifs mis en œuvre lors de la production et de la perception des voyelles du Français. Nos résultats démontrent l'existence de régions sensorimotrices de la voie dorsale impliquées dans le codage des buts phonémiques produits et perçus, régions dont les fonctions sont traditionnellement attribuées aux mécanismes de contrôle et de comparaison en ligne des représentations articulatoires et auditives. Ces résultats suggèrent ainsi l'existence de mécanismes perceptuo-moteurs communs impliqués dans le codage de voyelles perçues ou produites et résultats appuient ainsi l'hypothèse d'un possible couplage fonctionnel des systèmes de perception et de production de la parole au travers de représentations sensorimotrices distribuées des voyelles.

⁵⁷ "[...] it is entirely possible that RS, as indexed by fMRI adaptation paradigms, may reflect a combination of bottom-up automatic tuning of neuronal excitation and top-down building of perceptual expectation"

- 3^{ème} Partie -

Rôle fonctionnel de la voie dorsale
lors de la perception de la parole

Avant-propos

Les résultats observés lors des études IRMf présentées dans les Chapitres 1.2 et 2.2 démontrent l'activation commune des régions sensorimotrices de la voie dorsale lors de la perception et de la production des voyelles du Français. Cependant, ces résultats sont par nature corrélationnels et ne permettent pas de démontrer un rôle médiateur, fonctionnel, des régions activées.

La première étude présentée dans la troisième partie de cette thèse (Chapitre 3.1) avait pour but de déterminer l'implication fonctionnelle et causale des régions sensorielles et motrices de la voie dorsale lors de la perception de syllabes, par l'utilisation de la technique de stimulation magnétique transcrânienne et d'un paradigme conjoint d'adaptation auditive. Dans une seconde étude (Chapitre 3.2), nous avons utilisé une technique comportementale non invasive basée sur un entraînement moteur des articulateurs orofaciaux dans le but d'induire une recalibration/plasticité des représentations et processus moteurs possiblement impliqués lors d'une tâche d'identification auditive syllabique subséquente.

L'étude 3.1 a été réalisée en collaboration avec Marc Sato, Laurent Girin (Département Parole & Cognition, Laboratoire GIPSA-lab, UMR CNRS 5216 & Université de Grenoble), Pascale Tremblay (Faculté de Médecine, Département de réadaptation, Université Laval, Québec, Canada) et Vincent Gracco (Centre for Research on Language, Mind and Brain & School of Communication Sciences and Disorders, Université McGill, Montréal, Canada).

L'étude 3.2 a été réalisée en collaboration avec Marc Sato (Département Parole & Cognition, Laboratoire GIPSA-lab, UMR CNRS 5216 & Université de Grenoble), Arthur Glenberg (Department of Psychology, Arizona State University, Tempe, AZ, EU), Luigi Cattaneo (Center for Mind/Brain Sciences - CIMeC, University of Trento, Trento, Italie), Amélie Brisebois, Lucie Ménard (Département de Linguistique & Centre for Research on Language, Mind, and Brain, Université du Québec à Montréal, Montréal, Canada) et Anahita Basirat (INSERM-CEA Cognitive Neuroimaging Unit, Paris, France). L'élaboration du protocole, la passation des sujets et l'analyse des données de la partie expérimentale relative à un entraînement labial ont été effectuées dans le cadre d'un séjour de recherche d'Amélie Brisebois.

Ces études ont été principalement subventionnées par des financements de l'Institut National Polytechnique de Grenoble (Bonus Qualité Recherche "Modyc", "Modélisation dynamique de l'activité cérébrale") et du Plan Pluri-formations "Interactions Multimodales" de l'Université Stendhal (financement d'un séjour de trois mois au Centre for Research on Language, Mind and Brain de l'Université McGill pour la réalisation de l'étude 3.1).

Publications pertinentes

- Grabski K, Tremblay P, Gracco V, Girin L & Sato M. (soumis). A mediating role of the auditory dorsal pathway in selective adaptation to speech: a state-dependent transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*.
- Sato M, Grabski K, Glenberg A, Brisebois A, Basirat A, Ménard L & Cattaneo L. (2011). Articulatory bias in speech perception: evidence from use-induced motor plasticity. *Cortex*, 47(8):1001-3.

Chapitre 3.1

Implication de la voie dorsale dans les mécanismes d'adaptation sélective de la parole

Résumé / Abstract

En plus des traitements purement sensoriels, les théories constructivistes de perception de la parole proposent que les représentations articulatoires contraignent partiellement l'interprétation des entrées sensorielles à travers des mécanismes de simulation motrice interne. Cette approche constructiviste se trouve également au cœur des modèles neurobiologiques récents de la perception de la parole qui postulent l'existence d'une voie dorsale de traitement auditif dans l'hémisphère gauche, reliant les représentations auditives de parole au sein du cortex auditif temporal aux représentations motrices du système moteur frontal via le gyrus supramarginal et/ou la partie postérieure du gyrus temporal supérieur. La présente étude, conjuguant la technique de stimulation magnétique transcrânienne et un paradigme d'adaptation auditive, avait pour but de déterminer si la reconnaissance de la parole recrute effectivement la voie dorsale en examinant la contribution causale du cortex prémoteur ventral, du gyrus supramarginal et de la partie postérieure du gyrus temporal supérieur lors d'une tâche d'identification/catégorisation auditive syllabique. Dans ce but, les participants percevaient de manière passive une série de syllabes /ba/ avant d'effectuer une tâche de décision syllabique à choix forcé sur des syllabes ambiguës (situées à la frontière catégorielle entre /ba/ et /da/). En accord avec de précédentes études sur l'adaptation sélective en parole, un biais de réponse en faveur de la syllabe /da/ a été observé par comparaison à une condition sans adaptation auditive préalable. De manière cruciale, par rapport aux résultats observés sans stimulation, une impulsion unique de stimulation magnétique transcrânienne délivrée au début des cibles auditives syllabiques sur le cortex prémoteur ventral, le gyrus supramarginal ou la partie postérieure du gyrus temporal supérieur interagissait avec l'état initial de chacune des régions cérébrales stimulées et entraînait une augmentation de l'effet d'adaptation sélective. Ces résultats suggèrent une contribution de ces régions dans les mécanismes d'adaptation auditive en parole et démontrent le rôle médiateur de la voie dorsale dans les processus de catégorisation de la parole.

In addition to purely sensory processing, constructivist theories of speech perception argue that articulatory representations partly constrain phonetic interpretation of the sensory inputs through a process of internal motor simulation. The constructivist approach is also at the core of recent neurobiological models of speech perception postulating the existence of a left auditory dorsal processing stream, linking auditory speech representations in the auditory cortex with articulatory representations in the motor system, through sensorimotor interaction interfaced in the supramarginal gyrus and/or the posterior part of the superior temporal gyrus. The present state-dependent transcranial magnetic stimulation study aimed at determining whether speech recognition is indeed mediated by the dorsal pathway, by examining the causal contribution of the left ventral premotor cortex, supramarginal gyrus and posterior part of the superior temporal gyrus in an auditory syllable identification/categorization task. To this aim, participants listened to a

sequence of /ba/ syllables before undergoing a two forced-choice auditory syllable decision task on ambiguous syllables (ranging in the categorical boundary between /ba/ and /da/). In line with previous studies on selective adaptation to speech, responses were biased towards /da/ as compared to a control condition without prior auditory adaptation. Crucially, compared to the results observed without stimulation, single-pulse transcranial magnetic stimulation delivered at the onset of each target stimulus over the left ventral premotor cortex, the supramarginal gyrus or the posterior part of the superior temporal gyrus interacted with the initial state of each of the stimulated brain area by enhancing the adaptation effect. The results indicate that these regions contribute to auditory speech adaptation and demonstrate a mediating role of the dorsal pathway in speech categorization.

- *Mots clés : Perception et production de la parole, interactions sensorimotrices, voie dorsale, adaptation sélective, stimulation magnétique transcrânienne.*
- *Keywords: Speech perception and production, sensorimotor interactions, dorsal pathway, selective adaptation, transcranial magnetic stimulation.*

3.1.1 Introduction

Comment parvient-on à traiter le signal de parole pour extraire les invariants phonologiques et décoder l'information linguistique ? Un problème fondamental et ancien en parole concerne la difficulté à décrire de manière simple les relations existantes entre sons et phonèmes (voir par exemple Schwartz, Sato et Fadiga, 2008, 2011; Schwartz et al., 2012). Du fait de la grande variabilité du signal acoustique de parole, les processus de décodage acoustico-phonétique des propriétés et informations issues du signal acoustique restent en effet à ce jour encore largement incompris. Cette variabilité du signal acoustique de parole provient principalement de deux causes sous-jacentes et imbriquées. Tout d'abord, la production de tout message linguistique complexe implique des mécanismes de coarticulation entre phonèmes produits, soit une combinaison de gestes articulatoires à débit temporel élevé et largement parallélisée. Cette combinaison de gestes articulatoires est transformée de manière fortement non linéaire en une séquence de sons dans lesquels la correspondance entre sons et gestes est dès lors peu transparente (voir par exemple, Perkell et Klatt, 1986).

Bien que la perception de la parole soit parfois considérée comme un processus purement sensoriel, certaines théories de la perception de la parole proposent que l'interprétation phonétique des informations issues du signal de parole dépende également des connaissances procédurales articulatoires de l'auditeur (Liberman et al., 1967; Stevens et Halle, 1967; Liberman et Mattingly 1985; Liberman et Whalen, 2000; Schwartz et al., 2002, 2012; pour des revues récentes, Galantucci, Fowler et Turvey, 2006; Schwartz, Sato et Fadiga, 2008). L'implication du système moteur lors de la perception de parole a été initialement proposée comme une possible solution au problème de la variabilité du signal acoustique de parole et des relations non-linéaires entre unités phonétiques et leurs réalisations acoustiques. En se basant sur une approche constructiviste de la perception (Helmholtz, 1867), ces théories postulent que le locuteur et l'auditeur partagent un répertoire commun de primitives motrices (principe de 'parité motrice', Liberman et Mattingly, 1985) et que la perception de la parole repose sur des mécanismes de simulation motrice interne, qui permettraient de raffiner et contraindre l'interprétation phonétique des entrées sensorielles. Par exemple, dans la Théorie Motrice de la Perception de la Parole (Motor Theory of Speech Perception, Liberman et al., 1967; Liberman et Mattingly 1985; Liberman et Whalen, 2000), la capacité humaine à percevoir les sons de parole est supposée dépendre d'un synthétiseur du conduit vocal dont le rôle est de tester (par des processus d'analyse par synthèse, Stevens et Halle, 1967) lequel des buts articulatoires, ou des commandes motrices invariantes, serait à l'origine du signal acoustique entendu. De manière similaire, la Théorie de la Perception pour le Contrôle de l'Action (Perception for Action Control

Theory, Schwartz et al., 2002, 2012; voir aussi Basirat, Schwartz et Sato, 2012) propose également que les connaissances procédurales implicites de la production de la parole aident à fournir une meilleure spécification des entrées sensorielles de la parole et, de là, contribuent à l'analyse de scène en parole notamment dans des conditions d'écoute dégradée.

Cette approche constructiviste est aussi au cœur de modèles neurobiologiques récents de perception de la parole (Callan et al., 2004; Wilson et Iacoboni, 2006; Hickok et Poeppel, 2007; Skipper et al., 2007; Poeppel, Idsardi et van Wassenhove, 2008; Rauschecker et Scott, 2009; Hickok, Houde et Rong, 2011; Rauschecker, 2011). Ces modèles postulent l'existence d'une voie de traitement dorsale principalement localisée dans l'hémisphère gauche (voie 'how'), permettant de relier les représentations auditives de la parole issues du cortex auditif et les représentations articulatoires localisées dans le cortex prémoteur ventral et la partie postérieure du gyrus frontal inférieur, via des mécanismes d'interfaçage sensorimoteur au niveau du gyrus supramarginal (Rauschecker et Scott, 2009; Rauschecker, 2011) et/ou de l'aire 'SPT' (une région cérébrale au sein du planum temporale située à la jonction temporo-pariétale, Hickok et Poeppel, 2007). Le rôle de cette voie auditive dorsale, permettant un appariement sensorimoteur des représentations auditives et motrices, serait de permettre une comparaison des prédictions motrices fonction des connaissances procédurales de l'auditeur et des informations sensorielles réelles afin de contraindre l'interprétation phonétique des informations issues du signal de parole. Par exemple, le modèle de Skipper et collaborateurs ('hypothesis-and-test', Skipper et al., 2007) propose que des hypothèses multisensorielles sur les informations phonétiques issues du signal acoustique et/ou visuel de parole soient initialement élaborées dans la partie postérieure du gyrus temporal supérieur gauche. Dans le cas de signaux ambigus de parole, une génération des commandes motrices correspondant à ces hypothèses phonétiques serait effectuée via l'implication de régions motrices frontales (partie postérieure du gyrus frontal inférieur, cortex prémoteur ventral et cortex moteur primaire). Ces commandes motrices seraient utilisées pour prédire les possibles conséquences acoustiques et somatosensorielles associées aux mouvements articulatoires correspondants par des mécanismes de copies d'efférence envoyés aussi bien à la partie postérieure du gyrus temporal supérieur qu'au cortex somatosensoriel. Ces conséquences sensorielles simulées permettraient de contraindre l'interprétation phonétique des signaux perçus. De manière similaire, Rauschecker et Scott (2009, voir aussi Rauschecker, 2011; Scott et Johnsrude, 2003) proposent que la perception de la parole repose sur des traitements acoustiques de bas niveau au sein du gyrus de Heschl puis implique des processus de décodage acoustico-phonétique dans la partie antérieure du gyrus temporal supérieur et des régions frontales inférieures (voie ventrale). Suite à ces processus, ces hypothèses phonétiques seraient appariées avec les représentations articulatoires correspondantes dans le cortex prémoteur. Comme dans le modèle de Skipper et collègues (2007), ces représentations motrices seraient utilisées pour prédire les possibles conséquences sensorielles associées aux mouvements articulatoires correspondants par des mécanismes de copies d'efférence envoyés aussi bien à la partie postérieure du gyrus temporal supérieur qu'au gyrus supramarginal. Ces conséquences sensorielles simulées seraient comparées aux informations sensorielles réelles (voie dorsale) et permettraient de désambigüiser l'information phonétique issue du signal de parole.

En accord avec ces modèles, comme indiqué précédemment (voir Chapitres 1.2 et 2.2), de nombreuses études démontrent l'activation des régions motrices lors de tâches de perception de la parole selon des modalités auditive, visuelle ou audio-visuelle (par exemple, Nishitani et Hari, 2002; Calvert et Campbell, 2003; Callan et al., 2003; Paulesu et al., 2003; Callan et al., 2004; Möttonen et al., 2004; Wilson et al., 2004; Ojanen et al., 2005; Pekkola et al., 2005; Skipper, Nusbaum et Small, 2005; Pulvermüller et al., 2006; Wilson et Iacoboni, 2006; Skipper et al., 2007; Callan et al., 2010; Tremblay et Small, 2011). De plus, l'hypothèse d'une implication du système moteur fonction de l'intelligibilité/ambiguïté des stimuli perçus est appuyée par des études IRMf montrant une activité des régions du système moteur inversement corrélée au degré d'intelligibilité de la parole perçue, comme lors de parole masquée ou bruitée (Binder et al. 2004; Zekveld et al., 2006), lors de l'identification des phonèmes non-natifs vs phonèmes natifs (p. ex. Callan et al., 2004; Wilson et

Iacoboni, 2006) ou lors de la perception audiovisuelle de signaux incongruents (Jones et Callan, 2003; Ojanen et al., 2005; Pekkola et al., 2006; Skipper et al., 2007). Pris ensemble, ces résultats fournissent ainsi des éléments en faveur d'un rôle fonctionnel des régions motrices dans la perception de la parole et soutiennent l'hypothèse de processus d'appariement et d'intégration sensorimoteurs notamment dans des conditions d'écoute dégradée.

Cadre de l'étude

Néanmoins, malgré toute leur importance, ces résultats sont intrinsèquement corrélationnels et ne permettent donc pas de déterminer un possible rôle causal de ces régions dans la perception/compréhension de la parole. Par rapport à cette question, certaines études TMS ont permis de démontrer des effets d'interférence lors de tâches d'identification ou de discrimination syllabiques dans le cas de stimuli auditifs ambigus, par exemple présentés dans du bruit blanc (Meister et al., 2007; d'Ausilio et al., 2009; Möttonen et Watkins, 2009).

La présente étude a pour but d'étendre ces résultats et de déterminer la contribution causale de certaines des régions cérébrales constitutives de la voie dorsale - le cortex prémoteur ventral, le gyrus supramarginal et la partie postérieure du gyrus temporal supérieur - dans une tâche d'identification/catégorisation auditive syllabique. Pour ce faire, nous avons utilisé la méthode de stimulation magnétique transcrânienne conjointement à un paradigme d'adaptation auditive. Ce protocole, dit 'state-dependent', est basé sur l'application d'une procédure d'adaptation perceptive préalablement à la stimulation d'une région cérébrale lors d'une tâche subséquente de catégorisation et sur l'hypothèse que les effets de la stimulation dépendent de l'état initial de la région cérébrale stimulée (voir Silvanto, Muggleton et Walsh, 2008; Cattaneo et Silvanto, 2008; Cattaneo et al., 2010; Cattaneo, Sandrini et Schwarzbach, 2010). Plus spécifiquement, des mécanismes d'adaptation perceptive sont dans un premier temps induits par la présentation répétée d'un stimulus (ce de manière comportementale et sans stimulation), avec pour but de modifier l'activité de régions cérébrales censées participer aux processus de catégorisation des stimuli perçus. Lors de la tâche subséquente d'identification/catégorisation, l'application d'une stimulation magnétique sur l'une de ces régions lors de la présentation d'un stimulus cible permet de déterminer une possible modulation de la réponse du sujet, par comparaison avec une condition sans stimulation. Une telle modulation suppose une implication de la région stimulée dans la tâche de catégorisation et, également, sa sensibilité/participation à l'effet d'adaptation préalable. Au contraire, si aucune modulation comportementale n'est observée, cela implique que les populations neuronales stimulées ne sont pas sensibles à l'effet d'adaptation. Ce paradigme conjoint d'adaptation et de stimulation magnétique transcrânienne a ainsi pour but de démontrer une contribution causale des régions stimulées dans des tâches de décision/catégorisation en fonction de l'adaptation auditive induite préalablement, mais également d'augmenter la sélectivité fonctionnelle et résolution de la stimulation (pour une revue, voir Silvanto, Muggleton et Walsh, 2008).

Un paradigme d'adaptation sélective de la parole a été utilisé dans la présente étude afin d'induire une modification perceptive de syllabes ambiguës. Ce mécanisme d'adaptation sélective fait référence à la présentation répétée d'un son de parole particulier et d'une diminution du nombre de réponses reportées pour ce stimulus lors d'une tâche subséquente d'identification. Par exemple, dans l'étude princeps d'Eimas et Corbit (1973), suite à la présentation répétée d'une syllabe /ba/ ou /p^ha/, les auditeurs devaient catégoriser des syllabes issues d'un continuum /ba/-/p^ha/. L'écoute répétée de la syllabe /ba/ a entraîné une diminution des réponses /ba/ par rapport aux réponses /p^ha/, par rapport à une tâche d'identification identique réalisée sans adaptation préalable. Inversement, une augmentation des réponses /ba/ a été observée suite à la présentation répétée de la syllabe /p^ha/. En utilisant le même paradigme, Cooper et Lauritsen (1974) ont montré par la suite que l'écoute répétée de la syllabe occlusive non voisée /pi/ avait pour effet un raccourcissement du délai de voisement lors d'une tâche subséquente de production de la même syllabe. Puisque ce paradigme d'adaptation sélective induit aussi bien des modifications perceptives que motrices, il est

dès lors possible que les processus d'adaptation sélective dépendent des systèmes de perception et de production de la parole, probablement au travers d'effets d'habituation/fatigue de populations neuronales impliquées dans la détection des traits phonétiques (voir Cooper, 1979 et Samuel, 1986).

Dans la présente étude, il était demandé aux participants d'écouter de manière répétée la syllabe /ba/ avant d'effectuer une tâche de décision syllabique auditive à choix forcé sur des syllabes ambiguës (non bruitées mais situées à la frontière catégorielle entre /ba/ et /da/). En regard des études précédentes sur l'adaptation sélective, cette procédure avait pour but d'induire des modifications perceptuo-motrices et un biais de réponse vers la syllabe /da/, par comparaison à une condition contrôle sans adaptation auditive préalable. Lors de la tâche d'identification syllabique, une stimulation était appliquée en début de syllabe sur la partie dorsale du cortex prémoteur ventral (svPMC), le gyrus supramarginal (SMG) ou la partie postérieure du gyrus temporal supérieur (pSTG). Par rapport à une condition sans stimulation, la comparaison des résultats observés avec ou sans stimulation devait permettre de déterminer la possible contribution de ces régions dans les processus d'adaptation sélective et de catégorisation/reconnaissance de la parole.

3.1.2 Méthode

Protocole

Participants

Douze volontaires adultes canadiens francophones natifs ont participé (9 femmes; moyenne d'âge : 27 ans, ± 2 ans). Tous les volontaires étaient droitiers (selon le test de latéralité manuelle Edinburgh Handedness Inventory; Oldfield, 1971), avaient une vision normale ou corrigée et n'avaient pas de troubles de parole, d'audition ou de motricité. Les participants ont été examinés pour des éventuels troubles neurologiques, psychiatriques, linguistiques ou autres troubles médicaux et contre-indications à l'IRM et à la TMS (Wasserman, 1998). Tous les participants ont donné au préalable leur consentement éclairé et ont reçu une compensation financière pour leur participation. Cette étude a reçu un avis favorable du Comité d'Éthique de la Faculté de Médecine de l'Université McGill.

Stimuli

Un ensemble de onze syllabes variant par rapport au lieu d'articulation sur un continuum /ba-/da/ a été synthétisé (c_{-5} à c_{+5}). Ces stimuli ont été créés en référence à une méthode de synthèse et à des valeurs acoustiques utilisées lors d'une précédente étude portant sur la perception catégorielle (Serniclaes et al., 2001). Plus spécifiquement, une somme des trois ondes sinusoïdales relatives aux fréquences des trois formants F_1 , F_2 et F_3 a été générée puis modulée au moyen d'un filtre peigne synchronisé avec la fréquence fondamentale par une fonction arbitraire exponentielle négative avec une constante de temps de 10ms. Les valeurs formantiques des syllabes à l'extrémité du continuum ont été choisies de manière à correspondre à des valeurs appropriées en français pour les syllabes /ba/ et /da/. La différence du lieu d'articulation entre les syllabes /ba/ et /da/ a été créée en modifiant l'onset des fréquences de transition initiales (FT2 et FT3), correspondant à celles des deuxième et troisième formants (F_2 et F_3) en parole naturelle. L'onset de la fréquence FT2 variait de 1150Hz à l'extrémité /ba/ jusqu'à 1450Hz à l'extrémité /da/ en dix intervalles égaux de 30Hz. L'onset de la fréquence FT3 variait de 2250Hz à l'extrémité /ba/ jusqu'à 2750Hz à l'extrémité /da/ en dix intervalles égaux de 50Hz. Les fréquences finales des transitions FT2 et FT3 ont été fixées à 1300Hz et 2500Hz respectivement (ce qui correspond aux fréquences d'onset pour le stimulus médian c_0). Pour chaque stimulus, la fréquence initiale du premier formant (F_1) était de 200Hz et la fréquence finale de 750Hz. Pour chaque stimulus, les amplitudes des ondes sinusoïdales F_1 , F_2 et F_3 variaient de manière linéaire de 0.1, 0.001, 0.0001 à 0.3, 0.15 et 0.2 pour simuler les contributions respectives

des formants dans l'énergie du signal de parole pour ce type de syllabes. Le début du voisement (Voice Onset Time, VOT) était de -55ms, la durée de toutes les périodes de transition des fréquences était de 30ms et la durée du segment vocalique stable était de 130ms. Lors du voisement négatif et des périodes de transition des fréquences, la fréquence fondamentale F_0 a été fixée à 100 Hz. Lors du segment vocalique stable, un contour descendant de F_0 quasi-linéaire a été appliqué (de 106Hz à 96Hz).

Procédure expérimentale

Lors de l'étude, les participants étaient assis face à un écran d'ordinateur situé à environ 50cm. Les stimuli acoustiques ont été présentés de manière diotique à un niveau sonore confortable via des écouteurs compatibles EEG/TMS. Le logiciel Presentation (Neurobehavioral Systems, Albany, CA) a été utilisé pour contrôler la présentation des stimuli visuels et auditifs ainsi que pour enregistrer les réponses des participants.

Préalablement aux différentes tâches d'identification syllabique, deux étapes préparatoires ont été effectuées et concernaient (a) une procédure de recalage anatomique de manière à établir une correspondance entre l'image anatomique de haute définition du sujet et le système de pilotage TMS lié à la position de la bobine (coil) par rapport à la tête du participant ainsi que (b) une procédure de détermination du seuil d'excitabilité motrice au repos du sujet (voir ci-dessous).

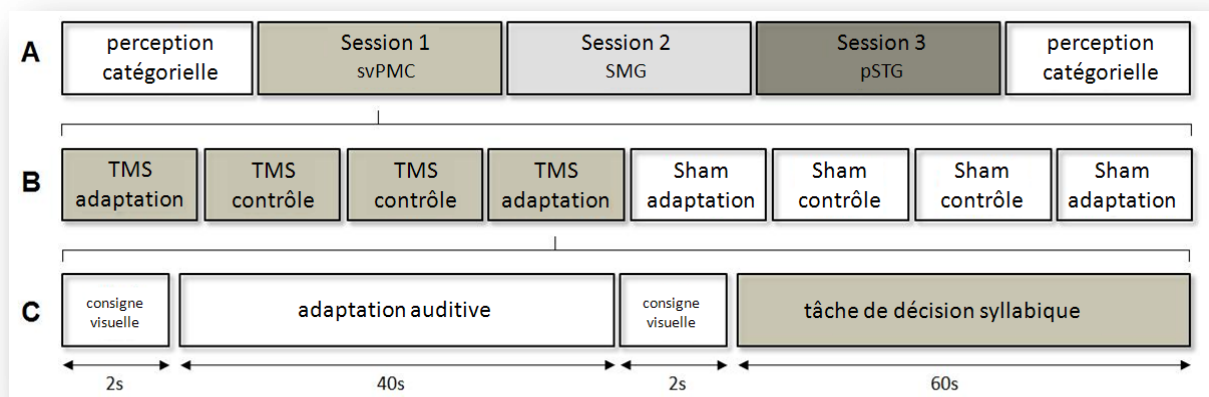


Figure 3.1 : Plan expérimental. A) L'expérience consistait en une tâche de perception catégorielle (sans stimulation), 3 tâches de décision syllabique avec stimulation des trois régions pSTG, SMG, svPMC et une seconde tâche de perception catégorielle (sans stimulation). B) Mise à part le site de stimulation, la procédure expérimentale était identique pour les 3 tâches de décision syllabique. Dans chaque tâche, 4 conditions ont été contrastées, en croisant le mode d'adaptation auditive (adaptation, sans adaptation) et le mode de stimulation (TMS, sans TMS ou sham). Chaque condition était présentée 2 fois au travers de 8 blocs expérimentaux. C) Chaque bloc expérimental consistait en une tâche d'adaptation ou une tâche contrôle de repos et une tâche de décision syllabique avec ou sans stimulation. Les ordres de passation relatifs aux sites de stimulation, au mode de stimulation et au mode d'adaptation ont été contrebalancés entre les participants et entre les 3 sessions.

L'expérience était constituée des différentes tâches suivantes (voir Figure 3.1):

- Une première tâche de perception catégorielle, réalisée sans stimulation, avait pour but de déterminer pour chaque participant les 3 syllabes (c., C₀, C₊) situées à la frontière catégorielle du continuum /ba/-/da/
- Trois sessions TMS (correspondant à chacun des trois sites de stimulation : pSTG, SMG, svPMC) avaient pour but de déterminer le rôle causal des régions de la voie dorsale dans les mécanismes d'adaptation sélective et de catégorisation de la parole.
- Une seconde tâche de perception catégorielle, réalisée sans stimulation, avait pour but de tester de possibles modifications à moyen-terme de la perception catégorielle du sujet.

Tâches de perception catégorielle : Dans le but de déterminer pour chaque participant les trois syllabes (c_- , c_0 , c_+) situées à la frontière catégorielle du continuum /ba/-/da/, une tâche de perception catégorielle a été réalisée sur les 11 stimuli du continuum (c_{-5} à c_{+5} ; voir Figure 3.2). Chaque essai débutait par une croix de fixation présentée au milieu de l'écran pendant 500 ms, suivie d'une présentation auditive d'une syllabe et se terminait par un écran vide pendant 1500 ms. Il était demandé aux participants d'identifier le plus rapidement et précisément possible la syllabe perçue, selon une procédure à choix forcé, en appuyant avec leur main droite l'une des deux touches du clavier correspondant à la syllabe /ba/ ou à la syllabe /da/. L'ordre des touches était contrebalancé entre les participants mais était identique pour toutes les tâches subséquentes. Chaque stimulus a été présenté 20 fois de manière pseudo-aléatoire (consistant en des présentations aléatoires des 11 syllabes par bloc) pour un total de 240 essais.

Une seconde tâche de perception catégorielle, identique en tous points à la première, a été réalisée après les sessions TMS, de manière à tester de possibles modifications à moyen-terme de la perception catégorielle du sujet

Bien que ces deux tâches de perception catégorielle aient été réalisées sans stimulation réelle, une impulsion magnétique était néanmoins délivrée 50ms avant le début du voisement de chaque syllabe avec la bobine éloignée de 10cm de la tête des participants de manière à reproduire l'artefact acoustique (clic) lié à chaque stimulation magnétique lors des sessions TMS.

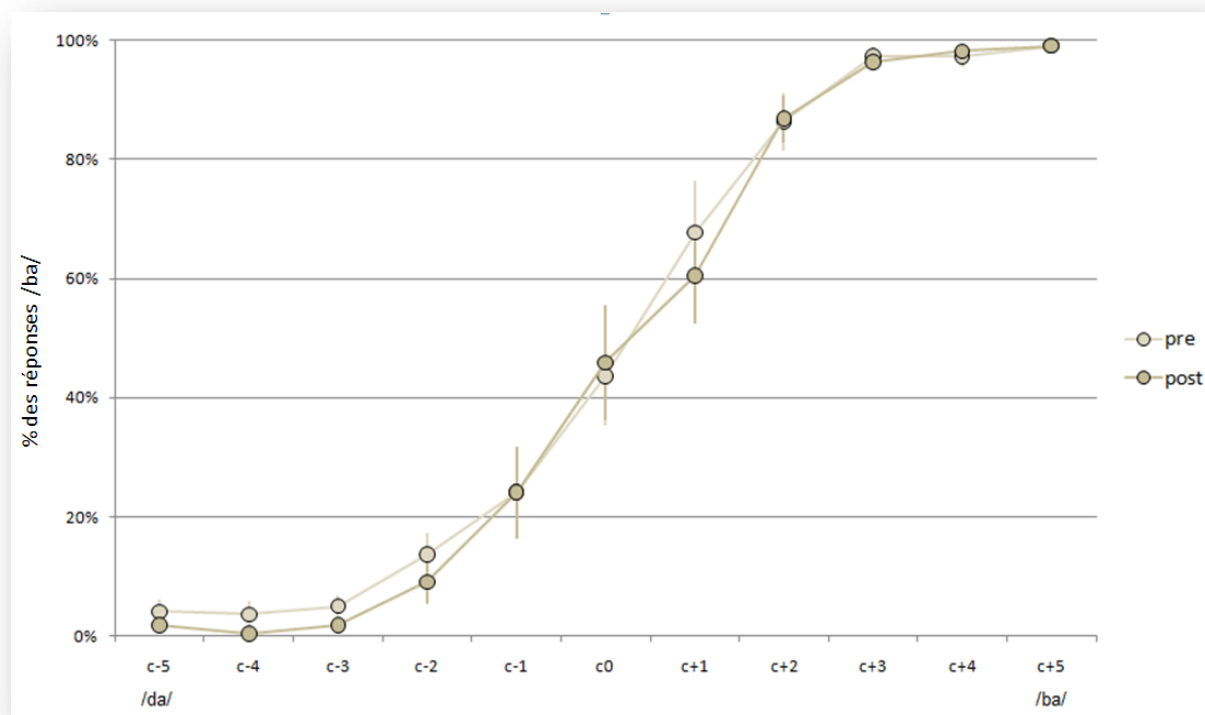


Figure 3.2 : Pourcentage moyen des réponses /ba/ en fonction des 11 stimuli du continuum /ba/-/da/ (c_{-5} à c_{+5}) observé lors des deux tâches de perception catégorielle précédant et suivant les sessions TMS. La première tâche de perception catégorielle permettait de déterminer pour chaque participant les 3 syllabes situées à la frontière catégorielle du continuum (c_- , c_0 , c_+).

Tâches de décision syllabique : Suite à la tâche de perception catégorielle initiale, trois sessions TMS ont été réalisées (une pour chaque site de stimulation : pSTG, SMG, svPMC). En dehors du site de stimulation, la procédure expérimentale était identique lors de ces trois tâches. Pour chacune d'elles, quatre conditions ont été contrastées, combinant le mode d'adaptation auditive (avec ou sans adaptation) et le mode de stimulation (avec ou sans stimulation). Chaque condition était présentée deux fois pour un total de huit blocs expérimentaux (voir Figure 3.1). Les ordres de passation relatifs aux sites de stimulation, au mode de stimulation et au mode d'adaptation ont été contrebalancés

entre les participants et entre les 3 sessions.

Chaque bloc consistait soit en une tâche d'adaptation auditive soit en une période de repos de 40s, toujours suivie d'une tâche de décision syllabique à choix forcé de 60s. Lors de la condition d'adaptation auditive, les participants écoutaient répétitivement la syllabe /ba/ (correspondant à la syllabe c_{+5} située à l'extrémité du continuum auditif) avant d'effectuer la tâche de décision syllabique. Cette syllabe était présentée 40 fois avec un intervalle de temps entre chaque présentation de 1s. Lors des 4 autres blocs, une condition contrôle correspondant à une période de repos de 40s, sans stimulation auditive, remplaçait la procédure d'adaptation. Suite à la condition d'adaptation ou de repos, une tâche de décision syllabique a été effectuée durant 60s relativement aux trois syllabes c_- , c_0 et c_+ situées à la frontière catégorielle du continuum /ba-/da/. Chaque essai débutait avec l'apparition d'une croix de fixation présentée au milieu de l'écran pendant 500 ms, suivie par la présentation d'une syllabe et se terminait par un écran vide pendant 1500 ms. Il était demandé aux participants d'identifier le plus rapidement et précisément possible la syllabe perçue, selon une procédure à choix forcé, en appuyant avec leur main droite l'une des deux touches du clavier correspondant à la syllabe /ba/ ou à la syllabe /da/. Pour chaque bloc, les 3 syllabes ont été présentées chacune 10 fois de manière pseudo-aléatoire (pas plus de deux répétitions de la même syllabe).

Dans quatre blocs consécutifs, une impulsion unique TMS a été délivrée relativement au site de stimulation 50ms avant le début du voisement de chaque syllabe présentée. Lors des quatre autres blocs, pour la condition sans stimulation (condition sham), une impulsion magnétique a également été délivrée 50ms avant l'onset de chaque syllabe mais avec la bobine éloignée de 10cm de la tête des participants de manière à reproduire l'artefact acoustique (clic) lié à chaque stimulation magnétique.

En résumé, pour chaque site de stimulation, chaque condition (TMS/adaptation, TMS/contrôle, sham/adaptation, sham/contrôle) impliquait 2 blocs initial ou final (bin1 et bin2) de 30 essais pour un total de 240 essais (dont 120 impulsions réellement appliquées).

Acquisition des données

Acquisition de l'image IRM anatomique et procédure de recalage anatomique

Une image structurelle du cerveau à haute résolution pondérée en T1 a été acquise préalablement à l'étude de manière à localiser précisément pour chaque participant les régions anatomiques devant être stimulées. Les images anatomiques ont été acquises avec un imageur corps entier de 1.5 Tesla (Siemens Sonata) à l'Institut Neurologique de Montréal. Les paramètres d'acquisition étaient les suivants : volume de 256x256x176mm, résolution isotrope de 1mm³, angle de bascule de 30°. Après acquisition, l'image anatomique a été incorporée au système Brainsight2 (Rogue Research, Montréal, Canada) de manière à établir une correspondance entre l'image anatomique de haute définition du sujet et le système de pilotage TMS lié à la position de la bobine (coil) par rapport à la tête du participant. Pour chaque participant, les sites de stimulation ont été identifiés sur l'image anatomique relativement aux repères suivants (voir Figure 3.3) :

- pSTG : Partie postérieure du gyrus temporal supérieur située ventralement par rapport à l'extrémité caudale de la branche postérieure de la fissure sylvienne (coordonnées moyennes x/y/z en mm dans l'espace MNI: -56 (±0.9) / -37 (±2.3) / 12 (±1.2).
- SMG : Partie la plus dorsale du gyrus supramarginal, située ventralement par rapport au sulcus intrapariétal (coordonnées moyennes x/y/z en mm dans l'espace MNI: -47(±1) / -49(±2.2) / 38 (±1.6).

- svPMC : Partie supérieure du cortex prémoteur ventral, région antérieure du gyrus précentral située à l'intersection du sulcus frontal inférieur et du sulcus précentral au niveau du gyrus postérieur précentral (coordonnées moyennes x/y/z en mm dans l'espace MNI: $-47 (\pm 0.8) / 5 (\pm 2) / 33 (\pm 1)$).

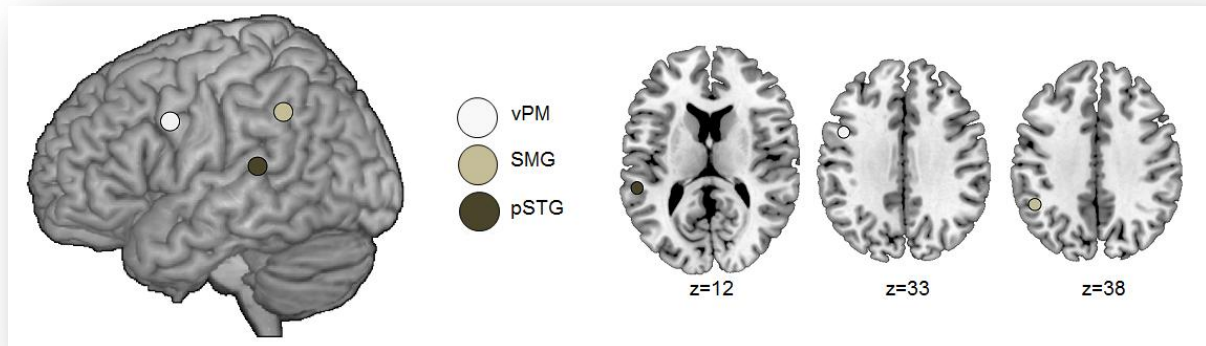


Figure 3.3 : Représentation sagittale (gauche) et axiale (droite) des coordonnées moyennes pour tous les participants des régions stimulées du cortex prémoteur ventral (svPMC), du gyrus supramarginal (SMG) et de la partie postérieure du gyrus temporal supérieur (pSTG).

Avant l'expérience, une procédure de recalage anatomique a été réalisée de manière à établir une correspondance entre l'image anatomique de haute définition du sujet et le système de pilotage TMS lié à la position de la bobine par rapport à la tête du participant. Pour ce faire, les positions de 4 points de repère (inion, nasion, les deux cartilages antérieurs de l'oreille) ont été repérées sur l'image anatomique et ensuite coalignées avec un système de traçage par infrarouge (tracking system Polaris, Northern Digital, Waterloo, Canada). Grâce à une procédure de normalisation spatiale, ces repères ont permis d'établir une correspondance spatiale et un suivi infrarouge de la position de la bobine par rapport à l'anatomie de chaque participant avec une précision de l'ordre d'un mm^3 et ainsi de déterminer les positions nécessaires pour la stimulation des 3 sites cérébraux.

Seuil d'excitabilité motrice

La TMS a été appliquée avec une bobine en forme de huit (height-coil ou butterfly coil) de 70 mm, connectée à un système de refroidissement par l'air et activée par un stimulateur magnétique à haute fréquence (Magstim Rapid 1400, Wales, U.K.). Afin de déterminer le seuil d'excitabilité motrice minimal de repos (RMT) pour chaque sujet, la bobine a été disposée sur la région du cortex moteur primaire de la main droite (le "hand knob" étant facilement repérable sur une image anatomique) et a été tenue de manière tangentielle par rapport au crâne, la poignée étant pointée dans un sens inféro-postérieur. La localisation de stimulation a été au fur et à mesure ajustée de manière à cibler l'aire de la main la plus excitable. Le RMT a été établi en tant que plus basse intensité de stimulation appliquée sur le cortex primaire de la main capable d'évoquer des contractions dans les muscles relâchés de la main droite pour au moins 5 sur 10 stimulations consécutives (Rossini et al., 1994).

Stimulation TMS

Lors de chaque tâche de décision syllabique, la bobine TMS a été d'abord placée sur le site de stimulation (soit pSTG, SMG ou svPMC) guidé par le système de localisation infrarouge. Une fois positionnée, elle a été tenue de manière tangentielle par rapport au crâne et ensuite sécurisée par un bras de positionnement. La bobine était orientée à 45° par rapport à l'axe rostro-caudal avec le manche pointant derrière pour la stimulation du svPMC et parallèlement à cet axe pour pSTG et SMG. Lors de l'expérience, la position de la bobine était contrôlée en ligne et ajustée après chaque mouvement de tête. Lors des conditions de stimulation (condition TMS), une impulsion unique a été appliquée 50ms avant le début du voisement de chaque syllabe avec une intensité de stimulation

égale à 110% du RMT (moyenne : $70\% \pm 2$ du maximum de stimulation). Lors de la condition sham, la bobine TMS a été éloignée de 10cm et tournée de manière perpendiculaire par rapport à la tête du sujet. Cette procédure permettait de reproduire l'artefact acoustique de stimulation sans pour autant stimuler magnétiquement la région.

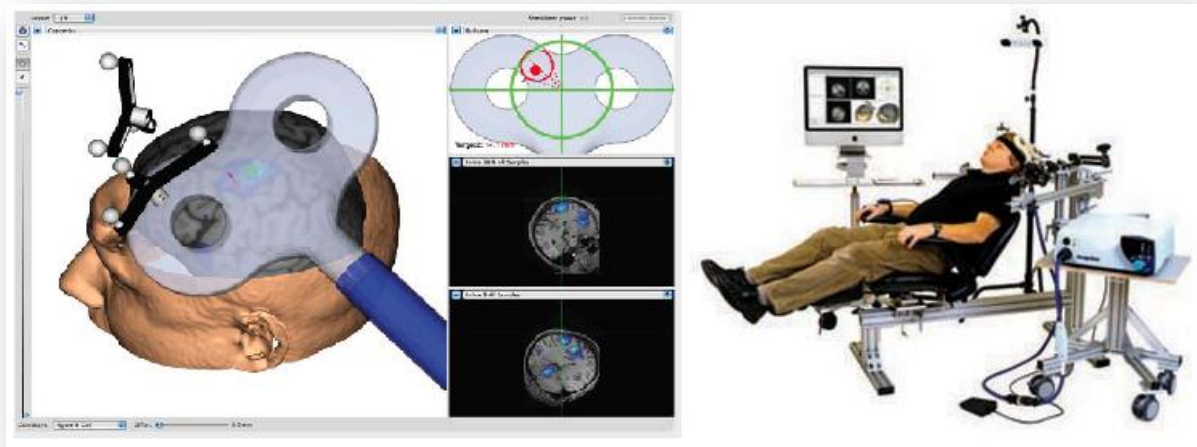


Figure 3.4 : Dispositif TMS. Gauche: représentation spatiale de la position du coil par rapport à la tête d'un participant et le site de stimulation. Droite : stimulateur magnétique et système de localisation de la bobine par infrarouge.

Analyse des données

Pour toutes les analyses suivantes, le niveau de significativité a été établi à $p < .05$. Lorsque nécessaire, des corrections Greenhouse-Geisser ont été appliquées. Les analyses post-hoc ont été réalisées à l'aide de tests de Newman-Keuls.

Tâche de perception catégorielle : La proportion des réponses /ba/ observées lors des deux tâches de perception catégorielle (avant et après les sessions TMS) a été calculée pour chaque participant et pour chacun des onze syllabes du continuum /ba/-/da/ (c_{-5} à c_{+5}). Une ANOVA à deux facteurs a été réalisée sur cette variable dépendante avec comme variables intra-sujet la syllabe (c_{-5} à c_{+5}) et la session expérimentale (avant et après)

Tâche de décision syllabique : La proportion des réponses /ba/ et le temps médian des réponses ont été calculés pour chaque participant, chaque mode de stimulation (TMS, sham), chaque site de stimulation (pSTG, SMG, svPMC), chaque mode d'adaptation (adaptation auditive, pas d'adaptation) et pour les deux bins (BIN1, BIN2). Une ANOVA à 4 facteurs a été effectuée sur ces variables dépendantes avec comme variables intra-sujet le mode de stimulation, le site de stimulation, le mode d'adaptation et le bin. Pour calculer spécifiquement le degré d'adaptation sélective en fonction des différentes conditions expérimentales, la proportion des réponses /ba/ et le temps médian des réponses observées pour les conditions avec adaptation auditive ont été soustraites de celles sans adaptation auditive préalable. Ces calculs ont été réalisés pour chaque participant, chaque mode de stimulation et chaque site de stimulation. Les réponses manquantes et celles pour lesquelles le temps de réponse était inférieur à 300ms ou supérieur 1500ms étaient considérées comme des erreurs et supprimées des analyses. Les résultats de deux participants n'ont pas été pris en compte dans ces analyses du fait d'un taux d'erreurs supérieur à 10% (moyenne : $2\% \pm 1$).

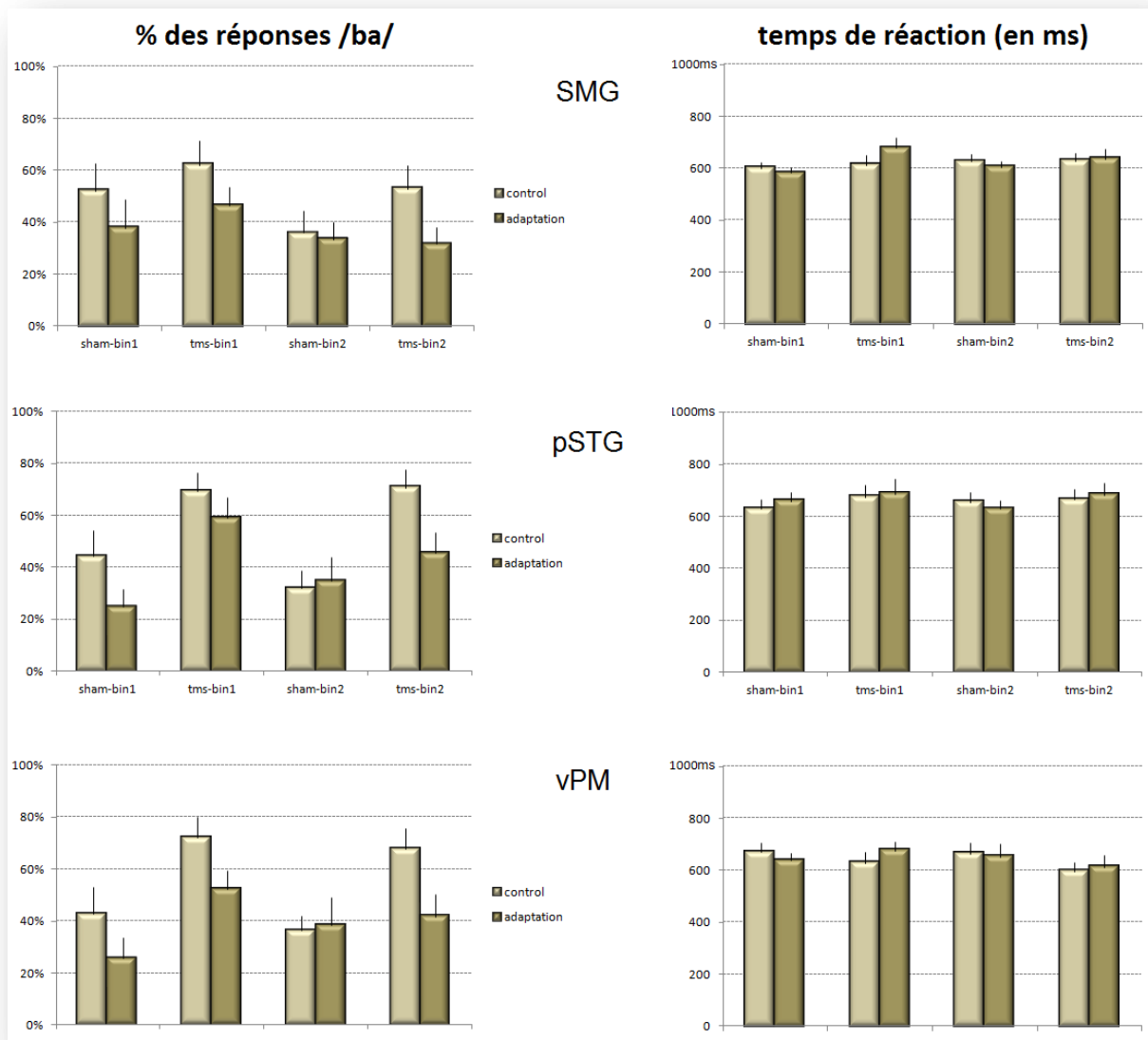


Figure 3.5 : Proportion moyenne des réponses /ba/ et temps de réponse médians observés lors des tâches de décision syllabique en fonction du site de stimulation (pSTG, SMG, svPMC), du mode de stimulation (TMS, sham), du mode d'adaptation (adaptation, contrôle) et du bin (BIN1, BIN2). Les barres d'erreur représentent les erreurs standards moyennes.

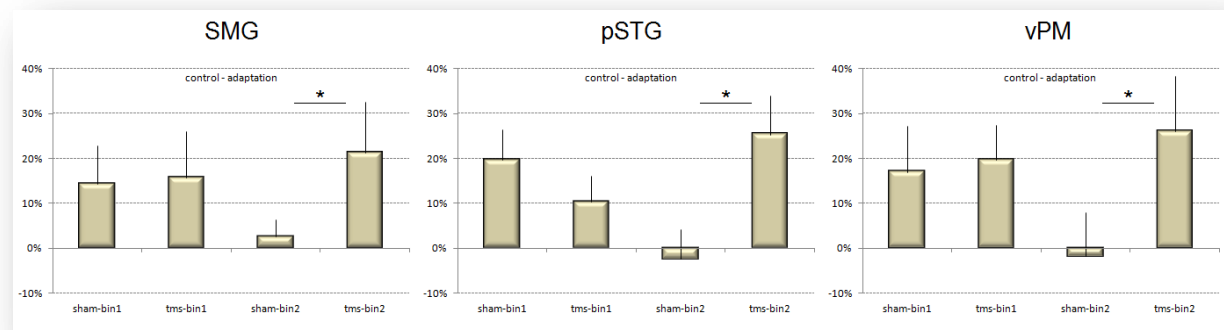


Figure 3.6 : Effet d'adaptation sélective (proportion des réponses /ba/ sans adaptation auditive préalable soustraite de celle observée avec adaptation auditive préalable) observée lors des tâches de décision syllabique en fonction des trois sites de stimulation (SMG, pSTG, svPMC), du mode de stimulation (TMS, sham) et du bin (BIN1, BIN2). Les barres d'erreur représentent les erreurs standards moyennes.

3.1.3 Résultats

Tâche de perception catégorielle

La Figure 3.2 illustre la proportion moyenne des réponses /ba/ en fonction des 11 syllabes du continuum auditif /ba-/da/ (stimuli c_{-5} à c_{+5}) observées lors des deux sessions de perception catégorielle (avant et après les sessions TMS). De manière attendue, la proportion moyenne des réponses /ba/ augmente au fur à mesure que le stimulus se déplace dans le continuum de l'extrémité /da/ vers l'extrémité /ba/ ($F_{(10,100)} = 137.7$, $p < .001$). Cependant, aucune modulation des résultats perceptifs et de la frontière catégorielle n'a été observée après les sessions TMS, ni l'effet de la session expérimentale (initiale, finale), ni l'interaction 'session x syllabe' n'étant significatifs. Ce dernier résultat indique que les procédures de stimulation et d'adaptation auditive lors des sessions TMS n'ont pas induit de modification à moyen-terme des représentations perceptives catégorielles des participants.

Tâche de décision syllabique

La Figure 3.5 montre la proportion moyenne des réponses /ba/ et les temps de réponse médian en fonction du site de stimulation (pSTG, SMG, svPMC), du mode de stimulation (TMS, sham), du mode d'adaptation (adaptation, contrôle) et du bin (BIN1, BIN2). La Figure 3.6 illustre le degré de l'effet d'adaptation sélective (proportion des réponses /ba/ observées sans adaptation auditive soustraite de celle observées avec adaptation auditive préalable) en fonction du site de stimulation (pSTG, SMG, svPMC), du mode de stimulation (TMS, sham) et du bin (BIN1, BIN2).

Scores perceptifs

Le score moyen global d'identification des syllabes perçues (c_{-} , c_0 , c_{+}) était de 46% ($\pm 3\%$) et similaire pour les 3 sessions TMS (en moyenne, 44% (± 6), 48% (± 3) et 47% (± 4) pour les sites SMG, pSTG et svPMC). Ce résultat confirme donc que les stimuli individuels c_{-} , c_0 et c_{+} étaient positionnés à la frontière catégorielle entre /ba/ et /da/ et ce pour chaque session TMS.

Une ANOVA à mesures répétées à 4 facteurs montre un important effet d'adaptation sélective ($(F_{(1,9)} = 17.2$, $p = .003$). Cet effet d'adaptation sélective, présent lors des trois sessions TMS (svPMC, SMG, pSTG) et pour les deux modes de stimulation (TMS, sham) correspond à la diminution prédite des réponses /ba/ dans la condition d'adaptation auditive par rapport à la condition contrôle (en moyenne, -9% et -19% pour les conditions sham/SMG et TMS/SMG, -9% et -18% pour les conditions sham/pSTG et TMS/pSTG, -8% et -23% pour les conditions sham/svPMC et TMS/svPMC).

Par ailleurs, une proportion accrue des réponses /ba/ a été observée lors des conditions TMS par rapport aux conditions sham correspondantes ($F_{(1,9)} = 8.4$, $p = .02$; en moyenne, +14% et +4% pour les conditions SMG/contrôle et SMG/adaptation, +32% et +23% pour les conditions pSTG/contrôle et pSTG/adaptation, +30% et +15% pour les conditions svPMC/contrôle et svPMC/adaptation). L'explication la plus plausible de cette augmentation des réponses /ba/ après stimulation pourrait provenir de l'artefact acoustique dû aux stimulations magnétiques, survenant 50ms avant le début du voisement des syllabes et pouvant partiellement masquer le bruit d'occlusion de la consonne. Bien que lors de la condition sham, l'artefact acoustique de stimulation ait été reproduit sans réelle stimulation (la bobine TMS étant éloignée de 10cm de la tête des sujets), cette hypothèse suggérerait que la conduction osseuse ait pu néanmoins biaiser les réponses des participants lors des conditions TMS (Nikouline, Ruohonen et Ilmoniemi, 1999).

En plus de ces effets, de manière cruciale, une interaction significative 'mode de stimulation x mode d'adaptation x bin' a été observée ($F_{(1,9)} = 6.5$, $p = .03$). Pour les syllabes présentées lors du premier bin, un effet d'adaptation apparaît aussi bien dans les conditions TMS que dans les conditions sham

(en moyenne, -14% et -16% des réponses /ba/ pour SMG/sham et SMG/TMS, -20% et -10% pour pSTG/sham et pSTG/TMS, -17% et -20% pour svPMC/sham et svPMC/TMS, aucune comparaison significative). En revanche, lors du second bin, l'effet d'adaptation disparaît pour les conditions sham mais persiste pour les conditions TMS (en moyenne, -3% et -21% des réponses /ba/ pour SMG/sham et SMG/TMS, 3% et -25% pour pSTG/sham et pSTG/TMS, 2% et -26% pour svPMC/sham et svPMC/TMS, toutes les comparaisons étant significatives). Il faut noter que cette interaction n'est pas due à l'artefact acoustique des stimulations. En effet, l'adaptation sélective (représentant les différences perceptives entre chaque condition d'adaptation et celle correspondante sans adaptation) a été calculée indépendamment pour les conditions TMS et sham (l'artefact acoustique était présent lors de toutes les conditions, notamment lors des conditions TMS/adaptation et TMS/contrôle). Aucun autre effet ou interaction n'a été observé.

Temps de réaction

Le temps de réaction moyen global était de 641ms (± 30) et similaire pour les trois sessions TMS (en moyenne, 621ms (± 26), 660ms (± 37) et 642ms (± 34) pour les sites SMG, pSTG et svPMC). L'interaction 'site de stimulation x mode de stimulation' est le seul effet significatif ($F_{(1,9)} = 5.29$, $p = .02$). Cette interaction est due à des réponses plus lentes après stimulation pour les sites SMG et pSTG, par rapport aux temps de réactions observés sans stimulation (en moyenne, -36ms et -34ms), alors que l'effet inverse est observé pour le site svPMC (+28ms). Aucun autre effet ou interaction n'a été observé.

3.1.4 Discussion

La présente étude, conjuguant stimulation magnétique transcrânienne et adaptation sélective, avait pour but de déterminer le possible rôle médiateur de la voie auditive dorsale dans la catégorisation et reconnaissance et ce, en examinant une possible contribution causale de régions svPMC, SMG et pSTG de l'hémisphère gauche lors d'une tâche de catégorisation syllabique auditive. Pour ce faire, des mécanismes d'adaptation perceptive ont d'abord été induits, par l'écoute répétée de la syllabe /ba/, dans le but de modifier l'activité des régions cérébrales de la voie dorsale censées participer aux processus de perception/catégorisation des stimuli perçus. De manière attendue, lors d'une tâche subséquente d'identification syllabique, un fort effet d'adaptation a été observé en comparaison d'une condition contrôle sans adaptation auditive préalable. De manière cruciale, par rapport aux résultats observés sans stimulation, une impulsion unique de stimulation magnétique transcrânienne délivrée au début des cibles auditives syllabiques sur le cortex prémoteur ventral, le gyrus supramarginal ou la partie postérieure du gyrus temporal supérieur interagissait avec l'état initial de chacune des régions cérébrales stimulées et entraînait une augmentation de l'effet d'adaptation sélective. Ces résultats suggèrent une contribution de ces régions dans les mécanismes d'adaptation auditive en parole et démontrent le rôle médiateur de la voie dorsale dans les processus de catégorisation de la parole.

Le paradigme 'state-dependent TMS' utilisé consistait en une procédure d'adaptation perceptive auditive préalablement à des stimulations magnétiques transcrâniennes délivrées lors d'une tâche de catégorisation syllabique subséquente. Ce paradigme, connu pour augmenter la sélectivité fonctionnelle et la résolution de la TMS (voir Silvanto, Muggleton et Walsh, 2008), a été utilisé afin de déterminer une possible contribution causale de trois régions cibles de la voie dorsale (svPMC, SMG et pSTG) dans les processus de décision/catégorisation syllabique et ce, en relation avec le stimulus induisant l'adaptation.

- Dans le but de déterminer individuellement les trois syllabes situées à la frontière catégorielle entre /ba/ et /da/, les participants ont tout d'abord effectué une tâche de perception catégorielle sans stimulation. Pour chaque participant, ces trois syllabes servaient de stimuli cibles lors des

trois sessions TMS suivantes. Lors de ces tâches, le score d'identification moyen de ces syllabes était proche du niveau du hasard (en moyenne, 44%, 48% et 47% de réponses /ba/ pour SMG, pSTG et svPMC), ce résultat confirmant par là même la bonne sélection de ces stimuli.

- En accord avec de précédentes études sur l'adaptation sélective en parole (Eimas et Corbit, 1973; pour une revue, voir Samuel, 1986), l'écoute répétée de la syllabe /ba/ induisait une forte adaptation sélective. Cet effet d'adaptation, correspondant à une diminution de réponses /ba/ lors de la tâche de catégorisation syllabique par comparaison à une tâche similaire mais réalisée sans adaptation auditive préalable, a été observé lors des trois sessions TMS (en moyenne, -14%, -13% et -15% de réponses /ba/ pour SMG, pSTG et svPMC). Il est également à noter qu'aucune modification perceptive n'a été observée entre les tâches de perception catégorielle initiale et finale, un résultat indiquant que les procédures de stimulation et d'adaptation auditive lors des sessions TMS n'ont pas induit de modification à moyen-terme des représentations perceptives catégorielles des participants.
- Lors des tâches de décision syllabique, une proportion accrue de réponses /ba/ a été observée après stimulation par comparaison avec une condition sham contrôle réalisée sans stimulation. Etant donné que l'artefact acoustique de stimulation était présent lors des conditions sham et contrôle, l'explication la plus possible à cet effet serait liée à des mécanismes additionnels de conduction osseuse lors de la condition TMS (la bobine étant placée en contact direct avec la tête du sujet). L'artefact acoustique de stimulation lors des conditions TMS, survenant 50ms avant le début du voisement des syllabes, aurait ainsi pu partiellement masquer le bruit d'occlusion de la consonne et, de là, biaiser les réponses des sujets vers la syllabe /ba/. En accord avec cette hypothèse, une étude conjuguant perception auditive, TMS et électroencéphalographie (Nikouline, Ruohonen et Ilmoniemi, 1999) a démontré que les potentiels évoqués auditifs liés à l'artefact acoustique des stimulations étaient affectés de manière importante par la position de la bobine, avec des amplitudes observées plus importantes lorsque celle-ci était appuyée directement contre le crâne du sujet par comparaison à une condition où elle était éloignée de quelques centimètres. Dans la précédente étude, il est donc possible que le bruit de stimulation propagé par conduction aérienne et osseuse lors des conditions TMS ait pu partiellement masquer l'occlusion consonantique et les fréquences de transitions des syllabes cibles (voir partie Méthodes) et ainsi biaiser les réponses des participants.
- De manière cruciale, par rapport à la condition sham sans stimulation, les impulsions délivrées lors du début de chaque stimulus cible ont interagi avec l'état initial des régions cérébrales stimulées, ce en augmentant l'effet d'adaptation. Lors des deux premiers blocs expérimentaux (BIN1), un effet d'adaptation fort a été observé pour les trois sites de stimulation aussi bien pour les conditions TMS que sham (en moyenne, -15%, -15%, -18% de réponses /ba/ pour SMG, pSTG et svPMC). Cependant, lors des deux derniers blocs (BIN2), alors que l'effet d'adaptation a presque entièrement disparu dans la condition sham (en moyenne, -3%, +3% et +2% de réponses /ba/ pour SMG, pSTG et svPMC), cet effet persiste et apparaît même renforcé dans la condition TMS (en moyenne, -21%, -25% et -26% des réponses /ba/ pour SMG, pSTG, et svPMC). Comme mentionné précédemment, il faut noter que cet effet ne peut pas être attribué aux artefacts acoustiques de stimulation étant donné que l'adaptation sélective a été calculée indépendamment pour les conditions TMS et sham et que l'artefact acoustique était présent lors de toutes les conditions (notamment lors des conditions TMS/adaptation et TMS/contrôle). Cette modulation de l'effet d'adaptation sélective induite par la TMS pour les 3 sites stimulés suggère donc l'existence de populations neuronales au sein de ces trois régions sensibles à l'adaptation auditive et jouant un rôle médiateur dans les processus de décision/catégorisation syllabique.

Cette modulation de l'effet d'adaptation sélective lors de la condition TMS pour les 3 sites stimulés de la voie dorsale apparaît en accord avec de précédentes études conjuguant la technique de TMS et une procédure d'adaptation ou d'amorçage visuel (par exemple, Silvanto, Muggleton et Walsh, 2008; Cattaneo et Silvanto, 2008; Cattaneo et al., 2010; Cattaneo, Sandrini et Schwarzbach, 2010; Cattaneo,

2010). Notamment, en étroite relation avec la présente étude, il a été montré que des effets d'adaptation dus à une exposition visuelle répétée d'actions transitives de la main ou du pied étaient modulés par la stimulation du cortex ventral prémoteur, du lobule pariétal inférieur et du sulcus temporal supérieur (Cattaneo, Sandrini et Schwarzbach, 2010; Cattaneo, 2010). De manière inverse, l'influence d'un entraînement moteur sur la perception a également été étudiée. Ainsi, des after-effects visuels, induits par l'exécution répétée d'actions transitives (les yeux bandés de manière à éviter tout retour visuel), étaient modulés par la stimulation du cortex ventral prémoteur gauche (Cattaneo et al., 2011; voir aussi Glenberg et al., 2010). Ces résultats apparaissent également en accord avec ceux obtenus dans de précédentes études IRMf où un paradigme d'adaptation a été utilisé afin d'étudier les substrats neuronaux du codage des buts d'action. Des effets de répétition suppression lors de l'observation répétée d'actions manuelles (correspondant à une baisse du signal BOLD; Grill-Spector et Malach, 2001; Grill-Spector, Henson et Martin, 2006; voir Chapitres 2.1 et 2.2) ont en effet été observés au niveau du cortex ventral prémoteur, du lobule pariétal inférieur et du sulcus intrapariétal adjacent dans l'hémisphère gauche, (Dinstein et al., 2007; Chong et al., 2008; Lingnau, Gesierich et Caramazza, 2008; Kilner et al., 2009). De plus, des paradigmes cross-modaux ont également montré des effets de répétition suppression aussi bien lorsque les actions manuelles étaient d'abord observées puis exécutées que lorsqu'elles étaient d'abord exécutées puis observées (Kilner et al., 2009; mais voir Lingnau, Gesierich et Caramazza, 2008). Ces résultats ont été largement discutés dans le cadre du rôle possible du système de neurones miroir chez l'homme dans le codage du but des actions perçus et produites (pour des revues, voir Rizzolatti, Fogassi et Gallese, 2001; Rizzolatti et Craighero, 2004).

Par rapport à la perception de la parole, notre étude confirme les résultats observés lors de précédentes études TMS et montrant qu'une perturbation temporaire de l'activité du cortex ventral prémoteur ou du cortex moteur primaire de l'hémisphère gauche induit une modulation de performance lors de tâches d'identification syllabique, en présence d'un bruit masquant (Meister et al., 2007; d'Ausilio et al., 2009) ou dans le cas de syllabes acoustiquement ambiguës (Möttönen et Watkins, 2009). De manière intéressante, une étude récente suggère également que des mécanismes de plasticité motrice, induits par un entraînement moteur et consistant en la réalisation répétée de mouvements linguaux ou labiaux, modifient les performances perceptives des sujets lors d'une tâche subséquente d'identification syllabique (Sato et al., 2011; voir Chapitre 3.2). De même, des travaux précédents ont démontré qu'une manipulation en ligne du feedback auditif lors d'une tâche de production de parole induit aussi bien des processus de compensation motrice articulatoire que des modifications perceptives (soit une recalibration des représentations sensorimotrices permettant de minimiser l'impact du feedback altéré; Shiller et al., 2009).

En accord avec ces travaux, nos résultats suggèrent que le cortex prémoteur ventral, le gyrus supramarginal et la partie postérieure du gyrus temporal supérieur jouent un rôle médiateur dans les processus de décision/catégorisation de la parole. Ce rôle des régions de la voie dorsale est en accord avec les théories constructivistes (Liberman et al., 1967; Stevens et Halle, 1967; Liberman et Mattingly 1985; Liberman et Whalen, 2000; Schwartz et al., 2002, 2012) et les modèles neurobiologiques de la perception de la parole (Callan et al., 2004; Wilson et Iacoboni, 2006; Hickok et Poeppel, 2007; Skipper et al., 2007; Poeppel, Idsardi et van Wassenhove, 2008; Rauschecker et Scott, 2009; Hickok, Houde et Rong, 2011; Rauschecker, 2011) postulant un rôle actif du système moteur dans la génération interne de candidats articulatoires permettant de contraindre l'interprétation et catégorisation phonétique du signal perçu par l'auditeur. Il est en effet proposé que les processus de décodage acoustico-phonétique des informations du signal de parole dépendent de projections bidirectionnelles entre systèmes sensoriels et moteur et de prédictions motrices issues des connaissances procédurales motrices de l'auditeur (notamment dans le cadre de copies d'efférences et de boucles sensorimotrices régulatrices, voir Wolpert, Ghahramani et Jordan, 1995; Kawato, 1999; Wolpert et Flanagan, 2001; Guenther, 2006; Hickok, Houde et Rong, 2011; Guenther et Vladusich, 2012; Perkell, 2012). Face à ces hypothèses, lors de la présente étude, l'augmentation de l'effet d'adaptation sélective suite à l'application d'une stimulation magnétique

transcrânienne sur le svPMC, le GSM et pSTG suggère ainsi une modification de l'état et excitabilité de populations neuronales au sein de ces régions induits par le paradigme d'adaptation sélective.

Enfin, bien que les résultats de notre étude suggèrent un rôle actif de la voie dorsale dans les processus de décision/catégorisation dans le cadre de stimuli ambigus, il est néanmoins important de rappeler que l'implication du système moteur lors de la perception de stimuli non bruités et non ambigus reste posée (par exemple, Hickok et Poeppel, 2007; Meister et al., 2007; Schwartz, Sato et Fadiga, 2008; d'Ausilio et al., 2009, 2012; Lotto, Hickok et Holt, 2009; Sato, Tremblay et Gracco, 2009; Scott, McGettigan et Eisner, 2009; Sato et al., 2011; Tremblay et Small, 2011). Comme indiqué précédemment, plusieurs études ne soutiennent pas l'idée d'un rôle médiateur du système moteur lors de conditions d'écoute optimales. Ainsi, des lésions au sein des régions motrices de la production de parole chez les patients souffrant de l'aphasie de Broca ne provoquent pas de déficits perceptifs systématiques, la plupart des patients aphasiques étant notamment capables d'effectuer des tâches de compréhension lexicale (Hickok et al., 2011). De même, des études de stimulation électrocorticale, lors d'opérations neurochirurgicales chez des patients éveillés, ou de stimulation magnétique transcrânienne impliquant une perturbation temporaire de l'activité de régions motrices ne montrent pas de déficits perceptifs concomitants dans la compréhension et reconnaissance de stimuli de parole lors de conditions d'écoute normale (Boatmann, 2004; Boatman et Miglioretti, 2005; Sato, Tremblay et Gracco, 2009; d'Ausilio et al., 2012; voir aussi Scott, McGettigan et Eisner, 2009). De futures recherches apparaissent donc nécessaires afin de tester plus avant une possible implication causale des régions de la voie dorsale lors de la perception de stimuli de parole non bruités et non ambigus.

3.1.5 Conclusion

La présente étude avait pour objectif de tester le possible rôle causal des régions de la 'voie dorsale' dans la catégorisation des sons de parole. Pour ce faire, nous avons effectué des stimulations magnétiques transcrâniennes sur le cortex prémoteur ventral supérieur, le gyrus temporal supérieur postérieur et le gyrus supramarginal antérieur afin d'induire une perturbation de l'activité de ces régions lors d'une tâche de catégorisation de syllabes. De plus, nous avons préalablement induit un effet d'adaptation auditive afin de moduler la perception catégorielle, de manière à augmenter la sélectivité fonctionnelle et la résolution spatiale des stimulations. Nos résultats démontrent l'impact de la stimulation magnétique transcrânienne conjointement à l'adaptation auditive par l'observation d'une augmentation de l'effet d'adaptation sélective après stimulation. Notre étude suggère ainsi une implication causale des régions de la 'voie dorsale' dans les mécanismes d'adaptation auditive et d'identification/catégorisation auditive syllabique.

Chapitre 3.2

Plasticité et implication du système moteur orofacial lors de la perception de la parole

Résumé / Abstract

Bien que de nombreuses études démontrent l'activation de régions motrices lors de la perception de la parole, un possible rôle médiateur du système moteur dans les traitements perceptifs des signaux de parole reste néanmoins débattu. Face à cette question, nous avons utilisé une technique comportementale non invasive basée sur un entraînement moteur des articulateurs orofaciaux dans le but d'induire une recalibration/plasticité des représentations et processus moteurs possiblement impliqués lors d'une tâche d'identification auditive syllabique subséquente. Plus spécifiquement, deux groupes de participants ont exécuté de manière répétée et silencieuse des mouvements linguaux ou labiaux pendant dix minutes de manière à induire une plasticité des régions impliquées dans le contrôle moteur de ces articulateurs. Suite à cet entraînement moteur, les participants ont effectué une tâche de décision syllabique à choix forcé entre les syllabes /pa/ et /ta/ (dont la production implique prioritairement la langue ou les lèvres, respectivement, pour leur articulation). Une analyse de détection du signal a été effectuée dans le but de séparer sensibilité auditive et possible biais de réponse. Par rapport à une tâche contrôle effectuée sans entraînement moteur préalable, aucun effet de l'entraînement sur la discriminabilité auditive n'a été observé. En revanche, des biais de réponse vers la syllabe /pa/ et vers la syllabe /ta/ ont été observés après l'entraînement moteur labial et lingual. Ces biais de réponse opposés après entraînement moteur labial et lingual suggèrent qu'une modification induite de l'activité du système moteur produit un biais perceptif spécifique et fonction de l'articulateur orofacial. Ces résultats suggèrent ainsi que l'entraînement moteur ne modifie pas la capacité de distinction auditive des deux syllabes et que le système moteur serait impliqué dans des processus de plus haut niveau de décision/catégorisation syllabique lors de la perception et reconnaissance de la parole.

Despite accumulating evidence that speech motor regions are activated in processing speech sounds, the question of whether articulatory processes mediate basic processes underlying speech perception remains debated. To test whether motor activity modulates performance in an auditory syllable identification task, we used a non-invasive behavioral technique based on activity-dependent plasticity in the orofacial motor system, with the goal of recalibrating action controllers that might be tapped by subsequent speech processing. Two groups of participants were required to repeatedly and silently perform either lip or tongue movements for ten minutes in order to induce changes in the corticomotor control of the orofacial musculature. Following the motor training, they underwent a forced-choice auditory syllable decision task between /pa/ and /ta/ syllables (which primarily involve lip or tongue movements when pronounced, respectively). A signal detection analysis was performed in order to separate auditory sensitivity to possible response bias. Compared to a control task performed without prior motor training, there were no effects of training on auditory discriminability. However, participants were more biased to respond /pa/ after the lip motor training,

while the opposite occurred after the tongue motor training. These opposite modulations of perceptual performance demonstrate that activity changes in the motor system, due to use-induced motor plasticity, specifically biased subsequent speech processing in an articulator-dependent manner. Altogether, these results suggest that motor training do not significantly impair the auditory ability to distinguish between the two syllables, but the motor system effect higher-level decision/categorization processes in auditory speech recognition.

- *Mots clés : Perception et production de la parole, interactions sensorimotrices, système moteur, voie dorsale, plasticité motrice.*
- *Keywords: Speech perception and production, sensorimotor interactions, motor system, dorsal pathway, motor plasticity.*

3.2.1 Introduction

Comme nous l'avons vu dans le précédent chapitre, de nombreux modèles et études soutiennent l'hypothèse selon laquelle l'interprétation phonétique des signaux de parole est déterminée ou tout du moins contrainte par les connaissances procédurales articulatoires de l'auditeur et la mise en œuvre de mécanismes de simulation sensorimotrice au sein de la voie dorsale. Ces modèles et études appuient ainsi l'idée d'un appariement automatique et spécifique des gestes articulatoires du locuteur jusqu'aux plans moteurs de l'auditeur lors de la perception de la parole.

Même pour les chercheurs opposés à une implication fonctionnelle du système moteur lors de la perception de la parole, l'idée d'un couplage des systèmes de perception et de production de la parole est largement admise (Hickok et Poeppel, 2000, 2004, 2007; Lotto, Hickok et Holt, 2009). Ce serait d'abord le cas chez l'enfant où une co-structuration des représentations sensorielles et motrices des unités de parole s'établirait de manière progressive au travers de mécanismes sensorimoteurs d'apprentissage orofacial et d'imitation vocale. Ces premières étapes exploratoires d'acquisition de la parole permettraient ainsi l'établissement d'associations entre buts moteurs et sensoriels, notamment auditifs (Guenther, 2006; Guenther et Vladusich, 2012). Chez l'adulte, la co-activation des régions auditives et motrices lors de la production des unités de parole pourrait entraîner la formation de liens sensorimoteurs par le biais de mécanismes d'apprentissage associatif et de principes hebbiens (Pulvermüller et al., 2006) et, de là, expliquer l'activation du système moteur observé lors de la perception de la parole sans rôle véritable dans les processus d'interprétation phonétique (Lotto, Hickok et Holt, 2009).

Néanmoins, nous l'avons vu, bien que les résultats de précédentes études démontrent un rôle actif du système moteur et des régions de la voie dorsale dans les processus de décision/catégorisation syllabique dans le cadre de stimuli ambigus, l'implication possible du système moteur lors de la perception de stimuli non bruités et non ambigus reste largement débattue. Pour rappel, il a été montré que des lésions frontales chez les patients aphasiques de Broca n'induisent pas de troubles spécifiques de compréhension lexicale, bien qu'une altération des performances de ces patients soit parfois observée lors de tâches méta-phonologiques (par exemple, Blumstein et al., 1977, Basso, Casati et Vignolo, 1977, Miceli et al., 1980; Caplan, Gow et Makris, 1995). Également, les résultats d'études de stimulation électrocorticale ou de stimulation magnétique transcrânienne montrent qu'une perturbation temporaire de l'activité de la partie operculaire du gyrus frontal inférieur gauche ou du cortex prémoteur ventral gauche entraîne une diminution de performance lors de tâches de d'identification/catégorisation syllabique uniquement en présence d'un bruit masquant (Meister et al., 2007; D'Ausilio et al., 2009), dans le cas des syllabes ambiguës du point de vue acoustique (Möttönen et Watkins, 2009) ou, enfin, lors du recrutement nécessaire et important de processus de segmentation et de mémoire de travail verbale (Boatmann, 2004; Nixon et al., 2004; Romero et al., 2006; Sato et al., 2009).

Cadre de l'étude

La présente étude avait pour but de déterminer si le système moteur articulatoire est impliqué fonctionnellement et spécifiquement lors de la perception de stimuli de parole bruités ou non et, si oui, si cette implication intervient dans des processus perceptifs de bas niveau ou au contraire lors de processus top-down de décision/catégorisation de plus haut niveau.

Pour ce faire, nous avons utilisé une technique comportementale non invasive basée sur un entraînement moteur des articulateurs orofaciaux. Plus spécifiquement, deux groupes de participants ont exécuté de manière répétée et silencieuse des mouvements linguaux ou labiaux pendant dix minutes de manière à induire une plasticité des régions impliquées dans le contrôle moteur de ces articulateurs. Suite à cet entraînement moteur, les participants ont effectué une tâche de décision syllabique à choix forcé entre les syllabes /pa/ et /ta/ (dont la production implique prioritairement la langue ou les lèvres, respectivement, pour leur articulation). Une analyse de détection du signal a été effectuée dans le but de séparer sensibilité auditive et possible biais de réponse impliqué dans des processus de décision de plus haut niveau. Puisqu'un tel entraînement orofacial à court terme a été associé à des mécanismes de neuroplasticité à court-terme et, notamment, une excitabilité neuronale accrue du cortex moteur primaire orofacial (Svensson et Romaniello, 2003; Boudreau et al., 2007; Sessle et al., 2007), une modulation spécifique de performance lors de la tâche de catégorisation syllabique après un tel entraînement moteur labial ou lingual démontrerait l'implication fonctionnelle du système moteur lors de la perception de la parole.

3.2.2 Méthode

Protocole

Participants

Trente six volontaires adultes de langue maternelle française ont participé à l'étude : 20 dans l'expérience 1 (entraînement moteur lingual; 8 femmes; moyenne d'âge : 29 ans (± 2)) et 16 dans l'expérience 2 (entraînement moteur labial; 6 femmes; moyenne d'âge 32 ans (± 3)). Tous les volontaires étaient droitiers (selon le test de latéralité manuelle Edinburgh Handedness Inventory; Oldfield, 1971), avaient une vision normale ou corrigée et n'avaient pas de troubles de parole, d'audition ou de motricité.

Stimuli

L'ensemble des stimuli consistait en 40 syllabes différentes : 10 /pa/, 10 /ta/, 10 /pa/ dans du bruit blanc (/pa/_{bruit}) et 10 /ta/ dans du bruit blanc (/ta/_{bruit}). Des occurrences multiples de syllabes /pa/ et /ta/ ont d'abord été enregistrées individuellement par 10 locuteurs français natifs dans une chambre sourde avec un taux d'échantillonnage de 44.1 kHz et une résolution de 16 bits. La production la plus clairement articulée de chaque syllabe a été sélectionnée pour chaque locuteur. Des procédures de normalisation de l'intensité (68dB) et de la durée (200ms, correspondant à un découpage de la syllabe 100ms avant l'occlusion consonantique et 100ms après) ont été réalisées en utilisant le logiciel Praat (Institute of Phonetic Sciences, University of Amsterdam, NL). Par la suite, un deuxième ensemble de syllabes masquées par du bruit a été créé. Pour chaque syllabe, un bruit blanc gaussien a été créé en relation avec le spectre des fréquences des syllabes /pa/ ou /ta/ et avec un délai progressif d'enveloppe du bruit de 5ms. Le masque durait 100 ms et était présenté simultanément au burst consonantique de la syllabe avec un ratio signal/bruit de -5dB.

Procédure expérimentale

Mise à part la nature de l'entraînement moteur (langue ou lèvres), la procédure expérimentale et le matériel utilisé étaient identiques dans les deux expériences. Les participants étaient assis dans une chambre sourde devant un écran d'ordinateur situé à une distance d'environ 50cm. Les stimuli acoustiques étaient présentés à un niveau sonore confortable au travers d'un casque audio. Lors de l'expérience, les participants ont réalisé deux tâches d'identification syllabique : l'une immédiatement après entraînement moteur (condition motrice) et l'autre sans entraînement moteur préalable (condition contrôle). L'intervalle entre les deux tâches était de 30 minutes et l'ordre d'entraînement était entièrement contrebalancé entre les participants dans les deux expériences.

Le logiciel Presentation (Neurobehavioral Systems, Albany, CA) a été utilisé pour contrôler la présentation des stimuli et pour enregistrer les réponses des participants. Lors de la tâche de décision syllabique, chaque essai débutait par une croix de fixation présentée au milieu de l'écran pendant 500 ms, suivie par la présentation d'une syllabe pendant 200 ms, et se terminait par un écran vide pendant 2500 ms. Il était demandé aux participants d'identifier le plus rapidement et précisément possible la syllabe perçue en appuyant avec leur main droite sur l'une des touches d'un clavier correspondant soit à la syllabe /pa/ soit à la syllabe /ta/. La désignation des touches de réponse a été contrebalancée entre les participants dans chaque expérience.

Avant l'expérience, un court entraînement était réalisé par les participants. Lors de la session expérimentale, la tâche de décision syllabique comportait 80 essais présentés de manière randomisée (20 /pa/, 20 /ta/, 20 /pa/_{bruit}, 20 /ta/_{bruit}).

Entraînement moteur

Procédure : Pour l'entraînement moteur des muscles linguaux, il était demandé aux participants de réaliser de manière répétée un mouvement vertical et d'appui de la langue contre la partie antérieure du palais dur, tout en gardant la bouche fermée afin d'éviter toute activité mandibulaire. Lors de l'entraînement moteur relatif aux muscles labiaux, les participants effectuaient des protrusions répétées des lèvres ('comme pour faire un bisou'). Les participants ont d'abord été entraînés à réaliser ces mouvements en contractant au maximum les muscles labiaux et linguaux correspondants, un retour visuel indiquant le niveau d'activité électromyographique de l'articulateur étant affiché sur un écran d'ordinateur. Ils étaient ensuite brièvement entraînés à effectuer de manière répétée ces mouvements linguaux ou labiaux (selon le groupe), avec pour instruction de maintenir une contraction musculaire pendant 2 secondes puis de relâcher les muscles les 2 secondes suivantes et ainsi de suite. Les mouvements étaient guidés par une croix de fixation apparaissant au milieu de l'écran toutes les 4s durant 2s.

Avant la tâche de décision syllabique, les participants effectuaient cet entraînement moteur pendant 10 minutes, soit 150 mouvements linguaux ou labiaux (150 x (2s de contraction + 2s de relaxation)). Il était demandé de garder une contraction musculaire stable (autour de 80% de la contraction maximale) tout au long de l'entraînement. Cette durée d'entraînement de 10 minutes a été choisie en se basant sur une étude précédente de stimulation magnétique transcrânienne montrant qu'un entraînement moteur lingual à court terme est associé avec un phénomène de neuroplasticité rapide au niveau du cortex moteur primaire humain au sein de l'aire orofaciale (reflété en terme d'augmentation des potentiels évoqués moteurs linguaux et d'une diminution du seuil d'excitabilité motrice au repos).

Electromyographie : Lors des entraînements moteurs, des enregistrements de l'activité électromyographique (EMG) du muscle mylo-hyoïdien (lors de l'expérience 1 avec entraînement moteur lingual) ou de l'orbicularis oris (lors de l'expérience 2 avec entraînement moteur labial) ont été acquis via le système EMG Biopac MP 150 contrôlé par le logiciel Acqknowledge (Biopac systems, Goleta, U.S.A.). Le signal EMG a été amplifié (x1000), échantillonné à 1Hz, filtré (filtrage passe-bande : 30-500 Hz) et stocké pour des analyses ultérieures. L'activité électromyographique de chaque

muscle a été enregistrée en utilisant des électrodes de surface Ag/AgCl (ELS254s, électrodes blindées avec un diamètre d'enregistrement de 4 mm) et un montage bipolaire. Lors de l'entraînement moteur de la langue, les deux électrodes ont été placées sous le menton à une distance respective de 5cm, symétriquement par rapport à la ligne médiane de la mandibule. Cet enregistrement devait permettre l'enregistrement de l'activité EMG du muscle mylo-hyoïdien (Cattaneo et al., 2007), responsable de l'élévation et l'abaissement de la langue lors de gestes de mastication et de déglutition (et dans le cas de la parole, pour produire des voyelles fermées et des consonnes vélaires; Epstein, Hacopian et Ladefoged, 2002). Le niveau de pression de la langue contre le palais coïncidant avec l'activité du muscle mylo-hyoïdien (Palmer et al., 2008), cet enregistrement fournissait ainsi une mesure indirecte des mouvements d'élévation de la langue lors de l'entraînement moteur. Pour l'entraînement moteur relatif aux muscles labiaux, l'activité musculaire labiale a été enregistrée au travers de deux électrodes éloignées à 2 cm l'une de l'autre au niveau du muscle orbicularis oris supérieur droit.

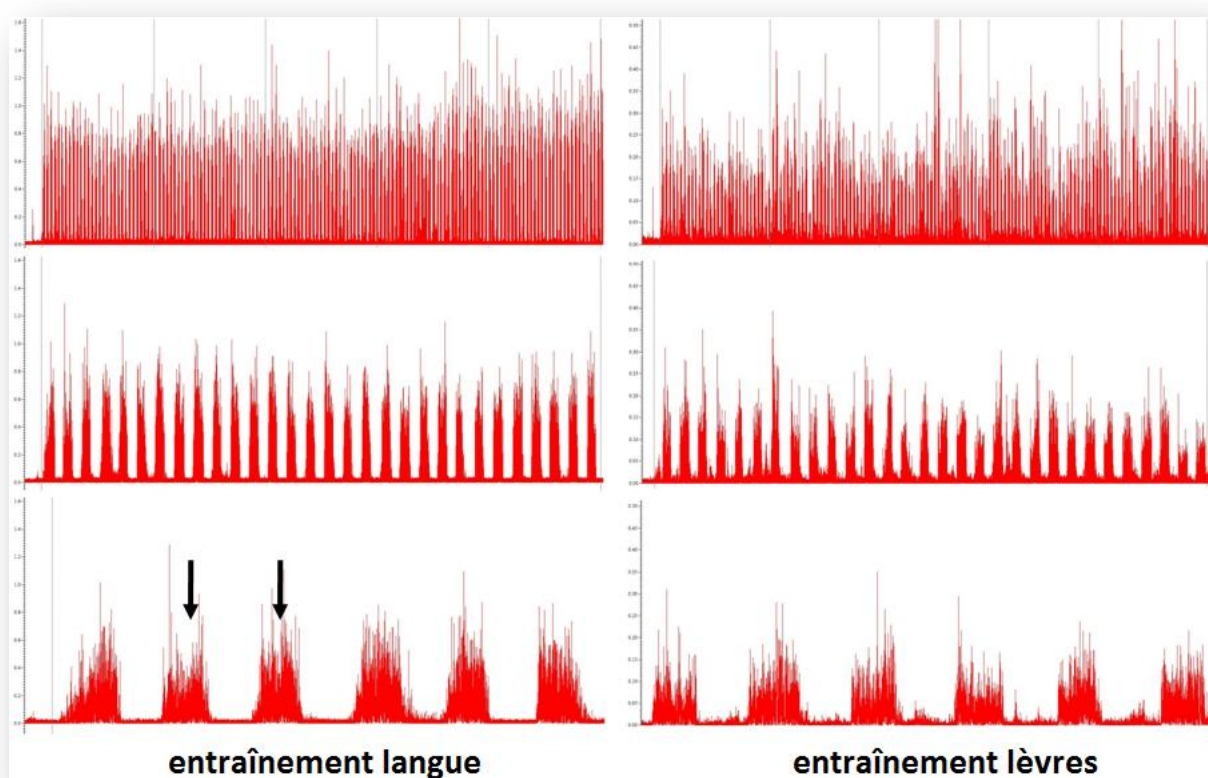


Figure 3.7 : Mesures de l'activité électromyographique rectifiée pour deux sujets du muscle mylo-hyoïdien lors de l'entraînement de la langue (expérience 1, à gauche) et du muscle orbicularis oris lors de l'entraînement des lèvres (expérience 2, à droite) et relative à 150 (haut), 30 (milieu) et 6 (bas) mouvements successifs (2s de contraction + 2s de relaxation).

Interrogés à la fin de l'expérience, les participants n'ont pas reporté de difficulté particulière à réaliser les entraînements moteurs. Du fait d'une faible qualité du signal EMG (mauvais contact des électrodes dû par exemple à une barbe ou une moustache), les données EMG de 6 participants lors de l'expérience 1 (langue) et de 5 participants lors de l'expérience 2 (lèvres) n'ont pu être analysées. Pour les autres participants, une inspection visuelle du signal EMG a confirmé la bonne réalisation des deux entraînements moteurs (voir Figure 3.7). Pour chaque participant, le signal EMG a été rectifié et les activités moyennes (aire sous la courbe) correspondant à 5 bins temporels consécutifs de 2 minutes ont été calculées. Afin de rendre possible une comparaison entre participants, ces valeurs ont été normalisées en pourcentages de la valeur moyenne d'activité observée sur tout l'entraînement (voir Figure 3.8). Pour chaque articulateur, une analyse de variance (ANOVA) a été

effectuée sur ces données avec comme facteur intra-sujet le bin (1 à 5). Lors des deux entraînements labial et lingual, un niveau constant d'activité EMG a été observé sans effet significatif du bin (entraînement lingual : $F_{(4,52)} = 1.75$; entraînement labial : $F_{(4,40)} = 0.50$).

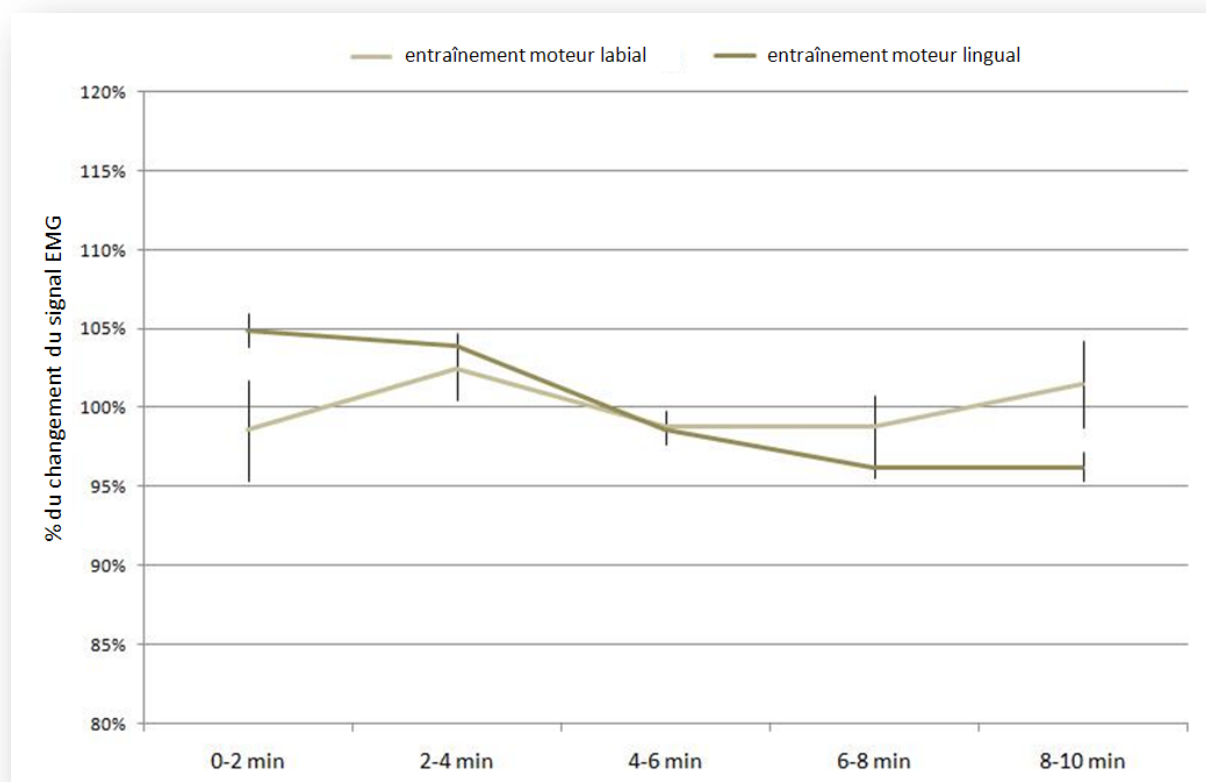


Figure 3.8 : Activités moyennes électromyographiques normalisées des muscles labiaux et linguaux lors des entraînements moteurs en fonction de 5 bins temporels de 2 minutes.

Analyse des données

Les pourcentages d'erreurs et les temps de réponse (TR) moyens observés pour tous les stimuli (/pa/, /ta/, /pa/_{bruit}, /ta/_{bruit}) ont été calculés pour chaque participant et chaque session (voir Figure 3.9).

Une ANOVA a été effectuée sur les valeurs moyennes de TR avec comme variables indépendantes intra-sujet la session expérimentale (entraînement moteur préalable, sans entraînement), le bruit (avec, sans) et la syllabe (/pa/, /ta/) et comme variable inter-sujets le type d'entraînement (lèvres, langue).

En ce qui concerne les scores perceptifs, nous nous sommes focalisés sur les paramètres de détection de signal, d' et β . Les mesures de d' indiquent la capacité auditive du sujet à distinguer les deux syllabes, indépendamment de tout biais possible de réponse. Une valeur de d' égale à 0 indique une totale incapacité de discrimination auditive, et d'autant plus large est la valeur de d' , d'autant plus grande est la capacité de discrimination. Les mesures de β indiquent dans le cas présent un biais de réponse vers la syllabe /pa/ (voir ci-dessous). Une valeur de β égale à 1 indique l'absence de biais. Plus le biais de réponse en faveur du /pa/ est important, plus la valeur de β décroît. Inversement, plus ce biais diminue, plus la valeur de β augmente. Pour chaque participant, chaque session et chaque condition, les valeurs de d' et de β ont été calculées en se basant sur les taux de réponses correctes et les taux de fausses alarmes par rapport à la syllabe /pa/ (voir Figure 3.9). Ainsi, répondre /pa/ lorsque la syllabe /pa/ est présentée est considéré comme une réponse correcte et lorsque la syllabe /ta/ est présentée comme une fausse alarme. Pour certains participants, le taux de

réponse correcte était de 100% (dans la condition non bruitée) et inversement le taux de fausse alarme était de 0%. Puisque ces cas ne permettent pas de calculer les valeurs de d' et de β , le taux de réponse correcte a été calculé comme le nombre de réponses correctes sur 21 essais (soit le nombre réel d'essais augmenté d'un essai) et le taux de fausse alarme a été calculé comme le nombre de fausses alarmes sur 21. Pour le calcul de d' et de β , les formules suivantes ont été utilisées (voir Macmillian et Creelman, 1991 et le site <http://psy2.ucsd.edu/~kang/sdt/sdt.htm>):

- $d' = \sqrt{2}[z(\text{taux de réponse correcte}) - z(\text{taux de fausse alarme})]$ où z représente le score standard
- $\beta = e^{(d' \cdot \sqrt{2} (z(\text{taux de réponse correcte}) + z(\text{taux de fausse alarme}))) / -2}$

Deux ANOVAs ont été réalisées relativement aux scores de d' et de β avec comme variables intra-sujet la session expérimentale (entraînement moteur préalable, sans entraînement), et le bruit (avec, sans) et comme variable inter-sujets le type d'entraînement (lèvres, langue).

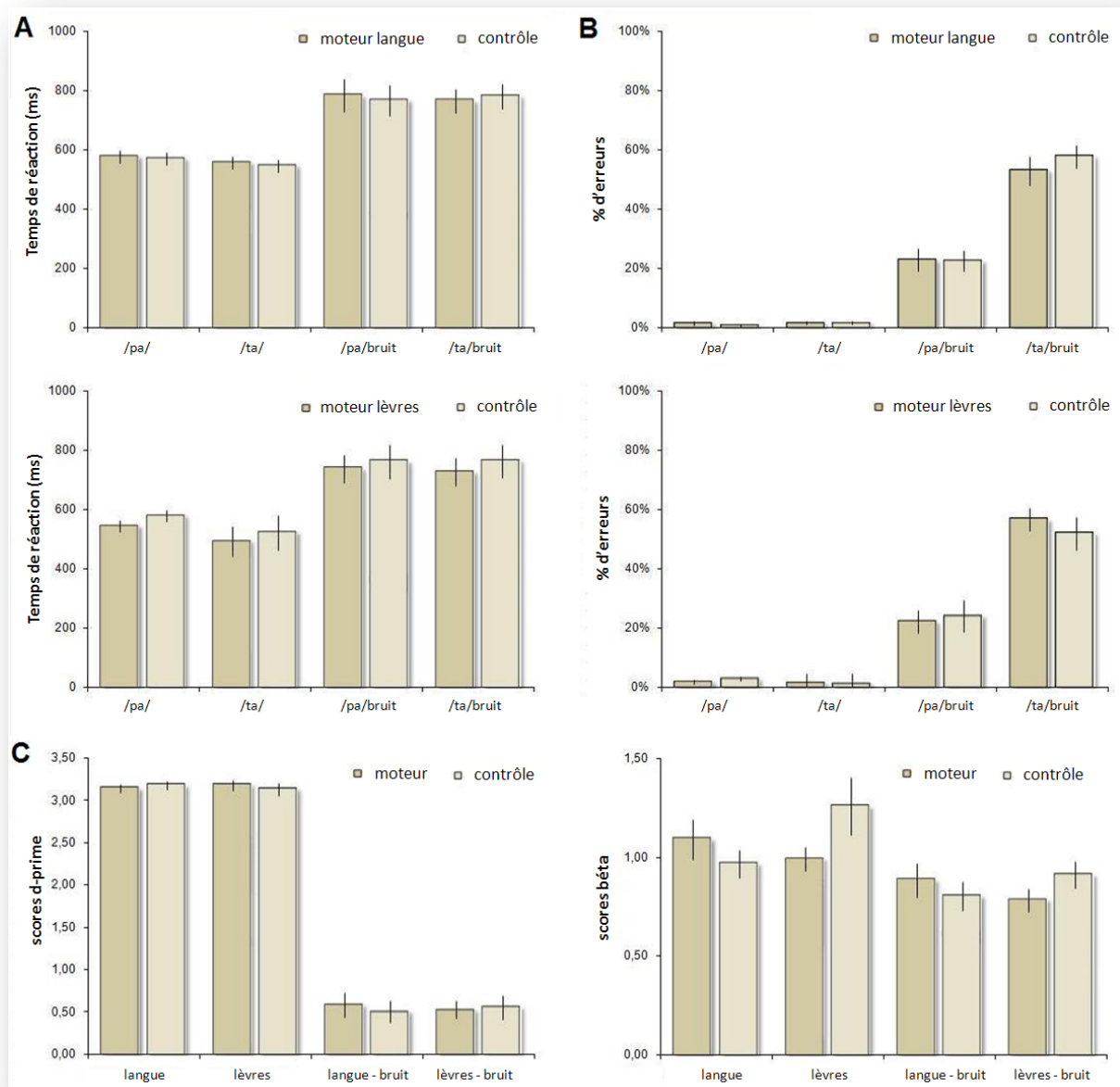


Figure 3.9 : Temps de réaction (A), pourcentages d'erreurs (B) et valeurs de d' et β (C) observées lors de la tâche de décision syllabique avec et sans entraînement moteur labial ou lingual préalable. Les barres d'erreurs représentent les erreurs standards.

3.2.3 Résultats

Scores perceptifs

L'ANOVA effectuée sur les valeurs de d' montre un large effet du bruit (les valeurs moyennes de d' pour les conditions avec et sans bruit étant de 0.54 et de 3.16; $F_{(1,34)} = 1278$, $p < .001$). Aucun autre effet ou interaction significatifs n'ont été observés, indiquant que ni l'entraînement moteur de la langue ni celui des lèvres n'a eu d'effet sur la capacité auditive à discriminer les deux syllabes.

L'ANOVA sur les valeurs de β a montré un effet significatif du bruit (les valeurs moyennes de β pour les conditions avec et sans bruit étant 0.85 et 1.08; $F_{(1,34)} = 13.98$, $p < .001$), ce qui indique que les participants étaient plus biaisés à répondre /pa/ lorsque la syllabe était présentée dans du bruit. De manière importante, une interaction significative entre la session expérimentale (avec ou sans entraînement moteur) et le type d'entraînement (lèvres, langue) a été observée ($F_{(1,34)} = 6.1$, $p < .02$). Après entraînement moteur des lèvres ($\beta = .89$), les participants étaient plus biaisés à répondre /pa/ que dans la condition contrôle ($\beta = 1.09$). Inversement, après entraînement moteur de la langue ($\beta = .99$), les participants étaient moins biaisés à répondre /pa/ (donc plus biaisés à répondre /ta/) que dans la condition contrôle ($\beta = .89$). Aucun autre effet ou interaction n'ont été observés.

Temps de réaction

De manière attendue, un large effet du bruit a été observé ($F_{(1,34)} = 72.14$, $p < .001$), avec des TRs plus courts pour les syllabes présentées sans bruit que celles présentées dans le bruit (en moyenne, 560ms vs 775ms). De plus, un effet significatif de la syllabe ($F_{(1,34)} = 6.23$, $p < .02$) et une interaction 'bruit x syllabe' ($F_{(1,34)} = 7.86$, $p < .01$) ont été observés. Ces effets étaient dus à des TRs plus longs pour les syllabes /pa/ que pour les syllabes /ta/ ($p < .001$; en moyenne, 580ms vs 540ms) présentées sans bruit et à l'absence de différence des TRs pour les syllabes /pa/ et /ta/ bruitées ($p = .84$; en moyenne, 776ms vs 774ms). Aucun autre effet ou interaction n'ont été observés, indiquant que ni l'entraînement moteur de la langue ni celui des lèvres n'ont eu d'effet sur les TRs.

3.2.4 Discussion

La présente étude avait pour but de tester un possible rôle fonctionnel du système moteur articulaire lors de la perception de la parole. Par rapport à une tâche contrôle effectuée sans entraînement moteur préalable, aucun effet de l'entraînement sur la discriminabilité auditive n'a été observé. En revanche, des biais de réponse vers la syllabe /pa/ ont été observés après l'entraînement moteur labial et, inversement, vers la syllabe /ta/ après entraînement lingual. Ces biais de réponse opposés après entraînement moteur labial et lingual suggèrent qu'une modification induite de l'activité du système moteur produit un biais perceptif spécifique fonction de l'articulateur orofacial. Ces résultats suggèrent ainsi que l'entraînement moteur ne modifie pas la capacité de distinction auditive des deux syllabes et que le système moteur serait impliqué dans des processus de plus haut niveau de décision/catégorisation syllabique lors de la perception et reconnaissance de la parole.

Une explication alternative peut néanmoins être envisagée. Il est en effet possible que les participants aient remarqué le rapport entre l'articulateur utilisé lors de l'entraînement moteur et la syllabe perçue correspondante et, de là, qu'ils aient utilisé stratégiquement cette relation pour modifier leurs réponses. Néanmoins, cette explication alternative apparaît peu probable pour deux raisons. La première est que la relation entre la tâche motrice et la réalisation articulaire des syllabes est loin d'être évidente, aucun des participants n'ayant d'ailleurs reporté avoir remarqué cette relation lors d'un débriefing faisant suite à l'expérience. La seconde est que l'effet observé

apparaît parfaitement cohérent avec de précédentes études TMS précédentes démontrant une modulation de performance lors de tâches d'identification/discrimination syllabique auditive suite à une stimulation du cortex moteur primaire orofacial ou du cortex prémoteur ventral de l'hémisphère gauche (Meister et al., 2007; D'Ausilio et al., 2009; Möttonen et Watkins, 2009; Sato, Tremblay et Gracco, 2009). Notamment, nos résultats sont en accord avec l'étude de d'Ausilio et collaborateurs (2009) montrant que la stimulation de l'aire de la langue ou des lèvres au sein du cortex moteur primaire gauche facilitait la reconnaissance de syllabes de manière spécifique du point de vue articulatoire. Cette double dissociation va dans le même sens que d'autres études TMS à impulsion unique et IRMf qui ont montré une somatotopie des activités motrices lors de la perception de la parole en relation avec les sons de parole perçus (Fadiga et al., 2002; Pulvermüller et al., 2006; Skipper et al., 2007; Roy et al., 2008; Sato et al., 2011). De notre point de vue, l'explication la plus plausible des résultats présents est celle d'une plasticité des représentations motrices suite à l'entraînement moteur et du rôle fonctionnel du système moteur articulatoire lors de la perception de la parole. Dans la mesure où l'entraînement moteur a induit une neuroplasticité à court-terme du système moteur orofacial lié au contrôle lingual ou labial (Svensson et Romaniello, 2003; Boudreau et al., 2007; Sessle et al., 2007), cette plasticité aurait pour conséquence de recalibrer l'activité du système moteur en relation avec l'articulateur utilisé. Ce mécanisme d'adaptation spécifique pourrait alors avoir affecté la reconnaissance des syllabes dans la tâche auditive suivante.

Bien que les données de notre étude vont à l'encontre de l'hypothèse selon laquelle la perception de la parole serait exclusivement dépendante de la récupération des gestes articulatoires à l'origine du signal perçu (Liberman et al., 1967; Liberman et Mattingly 1985; Liberman et Whalen, 2000), elles supportent clairement un rôle fonctionnel du système moteur. Nos résultats indiquent qu'une modulation à court-terme de l'état du système moteur n'a pas eu d'effet significatif sur la capacité auditive de distinction syllabique mais à un plus haut niveau de traitement lors de processus de décision/catégorisation. Ces résultats apparaissent ainsi cohérents avec certains modèles phonétiques, psycholinguistiques et neurobiologiques de la perception de la parole (Schwartz et al., 2002, 2012; Callan et al., 2004; Wilson et Iacoboni, 2006; Skipper et al., 2007; Poeppel, Idsardi et van Wassenhove, 2008; et Scott, 2009; Hickok, Houde et Rong, 2011; Rauschecker, 2011), qui postulent un rôle crucial du système auditif dans les processus de décodage acoustico-phonétique ainsi qu'un rôle actif du système moteur dans la génération interne de prédictions motrices utiles à une catégorisation phonétique finale.

Enfin, il est important de noter que l'absence d'interaction entre les facteurs 'bruit', 'session expérimentale' et 'entraînement' suggère indirectement l'implication du système moteur lors de la perception de la parole et ce, aussi bien pour des stimuli non bruités que pour des stimuli ambigus. Néanmoins, bien que non significatifs, les biais de réponses ont été plus importants en présence du bruit masquant. Ce qui va également dans le sens d'une activité du système moteur fonction de l'intelligibilité des entrées sensorielles, comme par exemple lors de l'identification auditive des phonèmes non-natifs par rapport à des phonèmes natifs (Callan et al., 2004; Wilson et Iacoboni, 2006), lors de la perception de parole masquée ou distordue (Binder et al. 2004; Zekveld et al., 2006) ou de la perception audiovisuelle de stimuli incongruents (Jones et Callan, 2003; Ojanen et al., 2005; Pekkola et al., 2006; Skipper et al., 2007). Enfin, les précédentes études TMS n'ayant pas reporté d'implication du système moteur lors de la perception de parole de stimuli non bruités (Sato, Tremblay et Gracco, 2009; d'Ausilio et al., 2011) n'ont pas utilisé une analyse de détection du signal, ce qui pourrait peut-être expliquer leurs conclusions négatives.

3.2.5 Conclusion

Lors de la présente étude, nous avons utilisé une technique comportementale non invasive basée sur un entraînement moteur des articulateurs orofaciaux dans le but d'induire une

recalibration/plasticité des représentations et processus moteurs possiblement impliqués lors d'une tâche d'identification auditive syllabique subséquente. Les biais de réponse opposés observés après entraînement moteur labial et lingual suggèrent qu'une modification induite de l'activité du système moteur produit un biais perceptif spécifique et fonction de l'articulateur orofacial. Ces résultats suggèrent ainsi que l'entraînement moteur ne modifie pas la capacité de distinction auditive des deux syllabes et que le système moteur serait impliqué dans des processus de plus haut niveau de décision/catégorisation syllabique lors de la perception et reconnaissance de la parole. Ces résultats apparaissent ainsi cohérents avec certains modèles neurobiologiques dans lesquels le système moteur permettrait de contraindre l'interprétation phonétique des signaux perçus.

- Discussion & Perspectives -

Discussion & Perspectives

Contrairement aux vues traditionnelles postulant une séparation stricte et fonctionnelle des systèmes de perception et de production de la parole, l'hypothèse principale défendue dans la présente thèse était celle, chez l'adulte, d'une coordination nécessaire des représentations auditives, somatosensorielles et motrices aussi bien pour la spécification, la production et la compréhension des unités de parole.

L'objectif premier de nos travaux était de déterminer un possible couplage fonctionnel des systèmes de perception et de production de la parole, par l'examen des activités cérébrales communes, sensorielles et motrices, lors de la production et de la perception des voyelles du Français. En regard de travaux récents portant sur des unités de parole plus complexes, l'étude des corrélats neuroanatomiques fonctionnels des voyelles du Français - considérées comme unités élémentaires de la parole, spécifiées et décrites en termes articulatoires et acoustiques précis autour d'un jeu universel de traits phonétiques, peu contaminées par des effets complexes de co-articulation et n'impliquant pas de traitements linguistiques morphologiques, sémantiques ou syntaxiques - devait permettre de préciser dans ce cadre minimal la possible nature sensorimotrice des représentations cérébrales de la parole.

Trois questions centrales et étroitement reliées l'une à l'autre ont ainsi guidé nos travaux :

1. Quelle est la nature - auditive, motrice, sensorimotrice - des représentations des unités de parole perçues et produites ? Existe-t-il un couplage fonctionnel des systèmes de perception et de production des voyelles du Français basé sur des connaissances procédurales motrices et sensorielles partagées par le locuteur et l'auditeur ?
2. Concernant l'implémentation neuronale de ces unités de parole, existe-t-il une topographie précise et spécifique, une phonémotopie, des voyelles ou des traits phonétiques associés ? Les représentations cérébrales des voyelles sont-elles distribuées au sein des régions auditives, somatosensorielles et motrices ?
3. Quelles sont les régions neurales sensibles à des mécanismes d'adaptation sensorimotrice et possiblement impliquées dans le codage du but des actions de parole produites ou perçues ?

Face à ces questions, nos travaux devaient en parallèle permettre de clarifier les structures cérébrales liées au contrôle moteur minimal orofacial lors de la réalisation de mouvements simples supralaryngés impliquant les mêmes articulateurs linguaux, labiaux et mandibulaires que lors de la production de la parole. Enfin, face à ce possible couplage fonctionnel des systèmes de perception et de production de la parole, une dernière série d'études avait pour objectif de déterminer une possible implication causale des régions sensorielles et motrices lors de la perception des unités de parole.

L'ensemble des travaux réalisés lors de la présente thèse apporte tout à la fois des éléments de réponse et de nouvelles pistes de recherche qui vont être discutés en quatre temps :

1. Quelles sont les régions neurales, communes et distinctes, impliquées dans le contrôle moteur orofacial des gestes supralaryngés, la production et la perception des voyelles du Français ?
2. Existe-t-il une cartographie cérébrale des gestes orofaciaux et des unités de parole perçues et produites ?
3. Quelles sont les régions cérébrales sensibles à des mécanismes d'adaptation à court-terme et au codage du but des actions orofaciales et vocaliques/phonémiques ?
4. Quel est le rôle fonctionnel des régions sensorielles et motrices de la voie dorsale dans la catégorisation des sons de parole ?

Réseaux du contrôle moteur orofacial, de la production et de la perception de voyelles

Les études IRMf 1.1 et 1.2 présentées dans le premier chapitre de cette thèse (et dans une moindre mesure les études 2.1 et 2.2 du second chapitre⁵⁸) avaient pour but de préciser les régions cérébrales activées lors de la réalisation de gestes articulatoires orofaciaux supralaryngés (lèvres, langue, mandibule) ainsi que lors de la production et de la perception de voyelles du Français (/i/, /y/, /u/, /e/, /ø/, /o/, /ɛ/, /œ/ et /ɔ/). Lors de ces études, nous avons utilisé la technique d'acquisition de sparse sampling dans le but de réduire le bruit du scanner lors des tâches de perception et de production des voyelles ainsi que d'éliminer de possibles artefacts liés aux mouvements articulatoires lors des tâches de production des mouvements supralaryngés et des voyelles. De plus, un enregistrement préalable de différentes occurrences de chaque type de voyelle (plus de 2000 occurrences) a été réalisé pour chaque sujet de manière à minimiser de possibles traitements de normalisation acoustique lors de la tâche de perception.

Nous rappellerons dans un premier temps les activations moyennes observées lors de ces trois tâches ainsi que les régions communes aux gestes supralaryngés et laryngés/vocaliques et aux systèmes de perception et de production des voyelles. Enfin, nous discuterons des limites et des perspectives de ces études.

Rappel des résultats

Le réseau minimal du contrôle moteur orofacial lié à la réalisation des gestes supralaryngés et laryngés/vocaliques (voir Figure D.1) implique des activations au niveau du sillon central et des gyri précentral et postcentral, avec deux larges clusters bilatéraux activés comprenant les portions supérieures des cortex prémoteur ventral, moteur et somatosensoriel primaires, ainsi que de l'aire motrice supplémentaire, du gyrus frontal inférieur, des ganglions de la base, de l'opercule pariétal et du gyrus supramarginal et, enfin, du cervelet supérieur. Ce réseau neural apparaît en accord avec de précédentes études IRMf portant sur le contrôle moteur orofacial lié à la réalisation de gestes articulatoires simples ou des actions orofaciales coordonnées plus complexes et permet ainsi d'affiner le réseau neuronal minimal impliqué dans le contrôle moteur de gestes orofaciaux simples, spécifique ou non à la production de la parole.

En accord avec de précédentes études sur le contrôle moteur de la parole, **le réseau de production des voyelles** (voir Figure D.1) révèle une ensemble de régions corticales et sous-corticales impliquées dans la préparation et l'exécution motrice, les traitements auditifs/phonologiques et dans des boucles sensorimotrices régulatrices. En plus de l'activation des cortex auditifs (régions auditives primaires, secondaires et associatives), le réseau global de production des voyelles implique ainsi des activations bilatérales des cortex prémoteur et de la partie operculaire du gyrus frontal inférieur postérieur ainsi que de l'insula gauche (programmation et coordination motrice), du cortex sensorimoteur primaire (pour l'exécution des commandes motrices par la voie pyramidale notamment), de l'aire motrice supplémentaire (initiation du mouvement) s'étendant jusqu'au cortex cingulaire antérieur (qui surveille des éventuelles erreurs), des lobules pariétaux supérieur (qui codent les représentations somatosensorielles du corps dans l'espace) et inférieur (codant les représentations phonologiques), du thalamus (relai principal des entrées et sorties entre le cortex et les autres régions du système nerveux), des ganglions de la base (noyau caudé et substance noire), le

⁵⁸ Du fait du paradigme de répétition suppression utilisé lors des études 2.1 et 2.2, nous concentrerons cette première discussion sur les résultats relatifs aux études présentées dans le premier chapitre. Nous rappelons cependant que malgré des activités moindres observées en regard de cette procédure d'adaptation, les résultats apparaissent tout à fait concordants entre ces différentes études.

cervelet (qui supervise la coordination motrice et la cohérence entre conséquences sensorielles des gestes prévus et leurs plans moteur) et du cortex visuel strié et extrastrié (qui correspondent à la perception visuelle des consignes).

Le réseau global de perception des voyelles (voir Figure D.2) implique de fortes activations auditives bilatérales au sein du gyrus/sulcus temporal supérieur (contenant les sites de traitements acoustiques primaires et secondaires et associatifs, impliqués dans les processus de décodage acoustico-phonétique) ainsi que des activations pariétales dans le gyrus supramarginal, le précuneus, l'opercule rolandique droit et des activations frontales dans la partie operculaire du gyrus frontal inférieur, s'étendant jusqu'au cortex prémoteur ventral adjacent⁵⁹. Cette activité des régions frontales et pariétales au sein de la voie dorsale lors de la perception des voyelles malgré un degré de simulation motrice supposé minimal (du fait du faible degré de complexité articulatoire et de co-articulation des voyelles ainsi que de la perception de sa propre voix qui minimise le recours pour le sujet à des processus de normalisation dans son propre espace acoustique afin d'inférer l'identité phonémique de la voyelle perçue) suggère ainsi l'existence d'interactions entre représentations auditives et motrices.

Contrôle moteur orofacial supralaryngé et production des voyelles (voir Figure D.1). Le réseau de production des voyelles englobe bien entendu le réseau de contrôle moteur orofacial minimal. Comme il est visible sur la Figure D.1 montrant la superposition des réseaux minimaux de contrôle moteur orofacial et de la production des voyelles, la production des voyelles implique en plus les activations auditives bilatérales du gyrus et sulcus temporal supérieur, des activations bilatérales au niveau du cortex moteur primaire dans la partie la plus ventrale (correspondant possiblement à l'aire de vocalisation ou à l'aire laryngée), et des activations plus amples au niveau du cortex prémoteur ventral s'étendant à la partie operculaire de l'aire de Broca (qui correspondrait aux représentations articulatoires et aux processus de planification articulatoire des sons de parole) et, enfin, deux clusters symétriques et proche de la scissure inter-hémisphérique (correspondant vraisemblablement aux muscles de la respiration)⁶⁰.

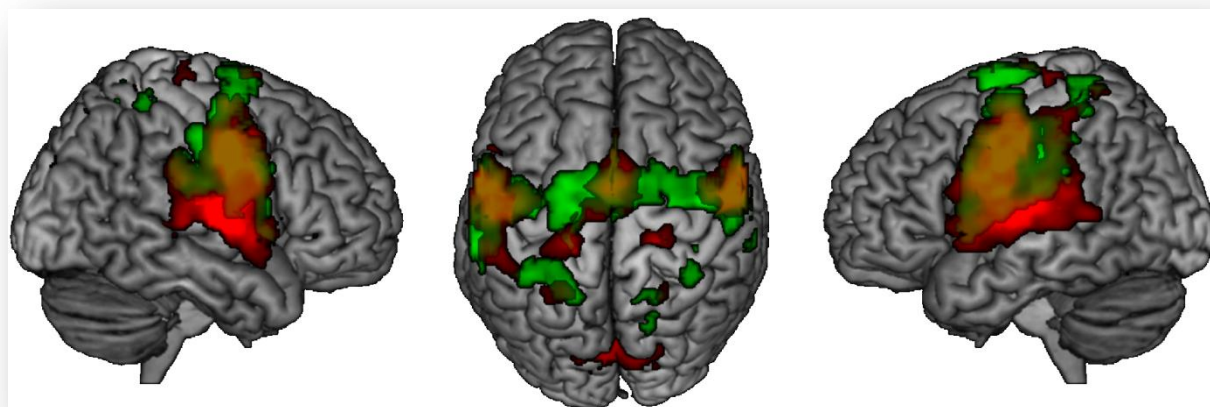


Figure D.1 : Superposition des régions cérébrales activées lors de la réalisation des mouvements orofaciaux supralaryngés (en vert) et de la production des voyelles (en rouge; seuil corrigé, $p < .05$ et taille minimale des clusters de 25 voxels).

⁵⁹ Des activations frontales supplémentaires ont été également observées dans la partie triangulaire du gyrus frontal inférieur gauche, dans le cortex dorsolatéral préfrontal gauche, le gyrus frontal supérieur gauche et moyen droit. Enfin, des activations ont également été observées dans les ganglions de la base (incluant le noyau caudé et la substance noire), dans les structures limbiques (incluant le thalamus, le cortex cingulaire et l'hippocampe), le noyau rouge et dans les régions visuelles striées et extrastriées (voir Discussion du Chapitre 1.2).

⁶⁰ Néanmoins, en dehors des activations temporales auditives pour la vocalisation de la voyelle /i/, les différences statistiques entre gestes supralaryngés et la voyelle /i/ sont restreintes à des régions ventro-latérales du cortex moteur et somatosensoriel primaire de l'hémisphère gauche, avec une intensité de signal plus ample pour la tâche de vocalisation (voir Discussion du Chapitre 1.1).

Perception et production des voyelles (voir Figure D.2). L'analyse de conjonction révèle des activations bilatérales auditives communes aux tâches de perception et de production des voyelles (incluant les cortex auditifs primaire, secondaire et associatif) et s'étendant caudalement à l'aire temporopolaire gauche, dorsalement au gyrus supramarginal gauche et à l'opercule pariétal droit. Les activations de la partie operculaire du gyrus frontal inférieur et du cortex prémoteur ventral adjacent sont également observées dans l'hémisphère gauche. Enfin, des activations additionnelles ont été observées dans le thalamus, la substance noire, le noyau rouge et le cortex visuel strié et extra strié. Ce résultat montre qu'en plus des aspects sensorimoteurs des représentations des voyelles aussi bien produites que perçues, ces deux systèmes de perception et de production partagent un réseau sensorimoteur au sein d'une voie de traitement postéro-dorsale dans l'hémisphère gauche qui englobe un ensemble de régions neurales liées au codage des représentations auditives, motrices ainsi qu'un site de convergence sensorimotrice (jonction temporo-pariétale) proposé comme lieu d'instanciation des représentations phonologiques dans de nombreux modèles et études (notamment en relation avec les processus de mémoire de travail verbale ou de traitements méta-phonologiques). L'apport principal de notre étude est ainsi de révéler les réseaux neuronaux sous-jacents aux voyelles perçues et produites, leur nature sensorimotrice et notamment de mettre en évidence un couplage fonctionnel des systèmes de perception et de production pour ces unités élémentaires de la parole.

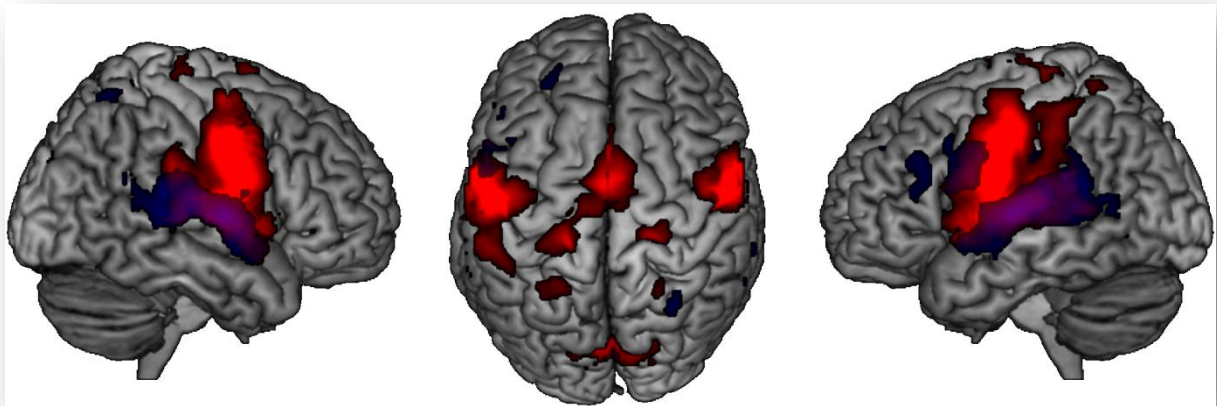


Figure D.2 : Superposition des régions cérébrales activées lors de la perception (en bleu) et de la production (en rouge) des voyelles (seuil corrigé, $p < .05$ et taille minimale des clusters de 25 voxels).

Limites et Perspectives

Certaines limites importantes liées aux protocoles expérimentaux doivent en premier lieu être soulignées :

- Le choix d'utiliser *a posteriori* la voyelle /i/ dans l'étude 1.1 est bien sûr contestable, il serait notamment plus judicieux d'utiliser une tâche plus spécifique au contrôle laryngé, comme par exemple un bourdonnement sourd ou un coup de glotte.
- D'autre part, les activations liées à la perception de voyelles lors de l'étude 1.2 sont comparées directement à la ligne de base qui consiste en une condition de repos. Une seconde condition contrôle consistant par exemple en la perception de stimuli acoustiques non-langagiers aurait permis de déterminer plus spécifiquement les traitements propres aux phonèmes perçus, en dehors des traitements auditifs de bas niveau ou non relatifs à la parole. L'utilisation d'une telle condition aurait été importante notamment en regard des débats concernant l'existence d'une voie auditive soit antérieure (Scott et Johnsrude, 2003; Rauschecker et Scott, 2009; Rauschecker, 2011) soit postérieure (Hickok et Poeppel, 2007) impliquée dans les processus de décodage acoustico-phonétique lors de la perception de la parole.

- Une autre limite à la tâche de perception concerne la perception par les sujets de leurs propres voyelles. La perception de sa propre voix pouvant impliquer des traitements auditifs et cognitifs particuliers, il aurait été intéressant de contraster la perception de ses propres voyelles avec la perception d'enregistrements d'autres locuteurs (en faisant varier le degré de correspondance avec l'espace vocalique du participant). Néanmoins, dans ce cas, il aurait été également important d'exclure les voyelles mi-ouvertes qui, du fait d'une plus grande variabilité inter-sujets, peuvent prêter à confusion en contexte isolé.

Certaines analyses additionnelles auraient pu également apporter des résultats complémentaires importants :

- Bien que nous ayons employé le terme d'"interactions" entre régions sensorielles et motrices en regard des résultats des études 1.1 à 2.2, nous n'avons pas réalisé d'analyses de connectivité fonctionnelle entre ces régions. La méthode d'acquisition en sparse-sampling (avec ici une acquisition intermittente toutes les 10 secondes) n'apparaît en effet pas adéquate à de telles analyses basées sur des corrélations temporelles d'activités. L'utilisation de courtes sessions d'enregistrement en continu suite à la perception ou la production d'une voyelle aurait pu permettre l'utilisation de telles analyses mais aurait considérablement allongé le temps des expériences. De même, une étude en électroencéphalographie ou en magnétoencéphalographie sur la perception des voyelles aurait pu apporter de précieuses informations sur le déroulement temporel des activités au sein de ces régions.
- En dehors de la somatotopie des articulateurs, la latéralisation des activités observées n'a pas été objet de l'étude de la présente thèse. Il est cependant envisageable de réaliser des analyses spécifique de latéralisation des réseaux impliqués dans la perception et la production des voyelles, tout comme dans la réalisation des gestes orofaciaux.

Enfin, il est important de mentionner certaines perspectives possibles de poursuite de ces travaux:

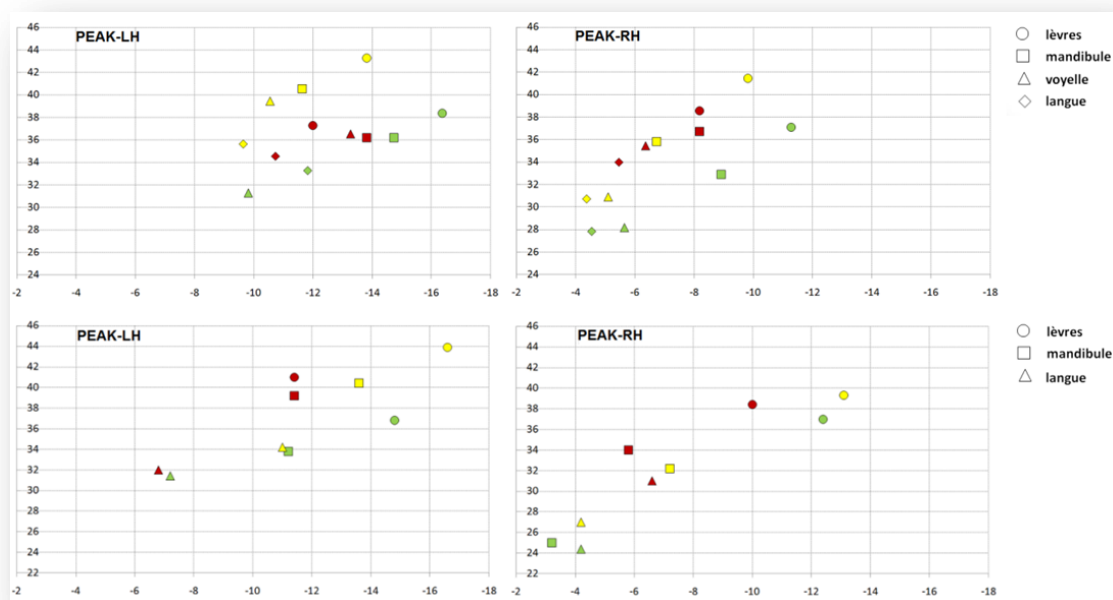
- En premier lieu, ces investigations pourraient être étendues aux consonnes, en vue d'acquérir une image plus complète des corrélats neuronaux liées aux phonèmes du Français. Dans une telle étude, il faudrait ainsi utiliser des syllabes simples de type consonne-voyelle. Dans ce cas, une tâche contrôle supplémentaire consisterait en la perception ou production de la voyelle impliquée. L'étude de ces syllabes serait intéressante puisque plusieurs études ont établi un lien entre la somatotopie des articulateurs supralaryngés contrôlant les lèvres et la langue et la production de syllabes impliquant une consonne initiale labiale ou linguale. D'autre part, il serait intéressant de comparer les réseaux observés pour les voyelles et les syllabes. En effet, il a été proposé dans certains modèles que les mécanismes de contrôle en ligne lors de la production de voyelles reposeraient principalement sur le retour auditif tandis que la production de consonnes nécessiterait des traitements additionnels basés sur les retours somatosensoriels (Hickok, 2012). Cette comparaison pourrait permettre de comparer la nature des boucles correctives dans les deux types de sons de parole, voyelles et consonnes.
- La présente étude nous a permis d'observer l'ensemble des activations corticales sous-jacentes à la production de voyelles mais non de distinguer les régions relatives aux systèmes de planification et d'exécution. Pour étudier cela, une future étude serait envisageable, dans laquelle pourrait être utilisé une tâche GO/NOGO ou des paradigmes similaires (comme par exemple dans Kell et al., 2010).
- Afin d'étudier l'impact de la maîtrise des actions vocaliques sur le réseaux observés, un projet postdoctoral récemment déposé a pour objectif de comparer les corrélats neuronaux des représentations de deux types d'actions vocales : des sons de parole et des vocalisations chantées et ce, chez les chanteurs (des participants experts dans les deux tâches) et des participants non experts, n'ayant pas de représentations particulières pour des actions vocaliques chantées.

Somatotopie et phonémotopie

Une question fondamentale reliée aux études présentées dans les deux premières parties de cette thèse concernait l'implémentation neuronale et une possible topographie précise et spécifique des gestes articulatoires supralaryngés ainsi que des voyelles perçues et produites ou des traits phonétiques associés. Nous rappellerons dans un premier temps les résultats observés lors des études 1.1 et 1.2 en regard de l'organisation somatotopique des gestes orofaciaux. Nous discuterons ensuite de l'absence de phonémotopie observée dans les études 1.2 et 2.2. Les limites et certaines perspectives à ces études seront enfin présentées.

Rappel des résultats

Représentation somatotopique des articulateurs orofaciaux : Les premiers travaux d'électrostimulation de Penfield et Rasmussen (1950) ont permis d'établir une organisation somatotopique des cortex primaires sensoriels et moteur pour les articulateurs de la parole, consistant en une organisation dorso-ventrale des représentations sensorimotrices des lèvres, du système de mastication, de la langue et du pharynx. Récemment, des études IRMf ont permis de confirmer cette topographie au sein du cortex sensorimoteur primaire pour les activations liées aux lèvres et à la langue (Lotze et al., 2000a; Hesselmann et al. 2004; Pulvermüller et al. 2006). L'emplacement de l'aire du larynx a été reportée par certains auteurs (Terumitsu et al., 2006; Brown, Ngan et Liotti, 2008) mais les conclusions restent débattues puisqu'il existe une discordance des résultats avec les travaux de Penfield et Rasmussen (1950) et le modèle de Guenther (2006). Les études 1.1 et 2.1 ont permis de confirmer et préciser cette somatotopie par l'observation d'une organisation dorso-ventrale des activités liées aux mouvements labiaux, mandibulaires, vocaliques/laryngés et linguaux au sein des cortex primaires moteur et somatosensoriels (voir Figure D.3).



En plus de la somatotopie dorso-ventrale lèvres/langue déjà démontrée par plusieurs études (et d'ailleurs la plus saillante dans nos résultats), l'apport de nos travaux est celui de placer la position de la mandibule entre ces deux articulateurs et ce, en accord avec la position du système de 'mastication' de Penfield et Rasmussen (1950). Par ailleurs, en vue des positions contradictoires de Steven Brown et collègues (2008), qui proposent l'emplacement du larynx et de la phonation de manière dorsale par rapport aux lèvres, et de Penfield et Rasmussen (1950) et Guenther (2006), qui proposent un emplacement ventral par rapport à la mandibule, nos études suggèrent que les représentations vocaliques se trouvent au même niveau que celles de la langue. Comme discuté précédemment, la vocalisation de la voyelle /i/ utilisée dans l'étude 1.1 implique non seulement le larynx mais aussi la langue. Il serait donc important dans une future étude d'utiliser des gestes laryngés le plus possible dépourvus d'implication d'autres muscles, comme par exemple des coups de glotte. Nos résultats suggèrent néanmoins une position laryngée/vocalique ventrale, au niveau de ou en dessous de la langue en accord avec la proposition de Guenther (2006).

Phonémotopie. Quelques études ont démontré un lien entre l'organisation somatotopique des articulateurs de la parole et le type de voyelles perçues de manière auditive, en montrant une organisation topographique des syllabes /pa/ et /ta/ correspondant aux aires des lèvres et de la langue, qui sont également les principaux articulateurs impliqués dans la production de ces deux types de sons de parole (par exemple, Pulvermüller et al., 2006). En plus d'une tonotopie des sons perçus et de cette 'somatotopie motrice perceptive' des articulateurs de la parole par un mécanisme de résonance motrice, certaines études ont également suggéré une organisation topographique pour les axes antéro-postérieur, latéro-médial et inféro-supérieur au sein du cortex auditif des différents traits phonétiques des voyelles (Shestakova et al., 2004; Obleser et al., 2003, 2006; Schäringer et al., 2011). Plus récemment, des études utilisant des analyses multi-variées de patterns d'activité (MVPA) ont permis de démontrer une organisation distincte des activités au sein du cortex auditif lors de la perception auditive de voyelles et de consonnes. Ainsi, Formisano et al. (2008) ont montré des patterns d'activité distincts pour les voyelles /i/, /a/, /u/ et Obleser et al. (2010) des patterns différents entre voyelles et syllabes de type consonne-voyelle, au sein du gyrus temporal supérieur. Dans le cadre d'une possible phonémotopie au sein des régions sensorielles et motrices, nous avons cherché des différences d'activations (ou de pics d'activation) pour différents traits phonétiques ou types de voyelles. Cependant, aucune différence statistiquement significative n'a pu être trouvée en regard des analyses univariées utilisées. Nos résultats ne permettent donc pas de démontrer l'existence d'une organisation phonémotopique des voyelles mais, face aux études précédentes et au couplage fonctionnel observé des systèmes auditifs et moteurs, suggèrent néanmoins l'existence d'un codage sensorimoteur distribué de ces voyelles.

Nos résultats permettent cependant de montrer des différences d'activation en fonction du degré d'ouverture des voyelles perçues et produites. Lors de la tâche de perception, une activité accrue a été observée au niveau du cortex auditif, du gyrus frontal inférieur, du gyrus cingulaire et de l'insula (voir Chapitre 2.2). Lors de la production des voyelles, une activité accrue pour ces voyelles a été observée au niveau de l'aire motrice supplémentaire et du gyrus cingulaire, du gyrus frontal inférieur, du cortex prémoteur et somatosensoriel, du gyrus supramarginal et du précuneus et enfin du cervelet. Nous interprétons ces résultats en terme de codage différencié selon le degré de maîtrise articulatoire et notamment, pour la tâche de production, d'une nécessité d'implication plus conséquente des régions motrices lorsque le son de parole n'est pas parfaitement maîtrisé, ce qui est le cas pour les voyelles mi-ouvertes en contexte isolé par rapport aux voyelles fermées et mi-fermées, que les locuteurs francophones ont l'habitude de produire isolément.

Limites et perspectives

- Comme expliqué précédemment, nous avons ajouté *a posteriori* la tâche de vocalisation de la voyelle /i/ afin de comparer les corrélats neuronaux des gestes laryngés aux gestes supralaryngés. Il serait important d'utiliser pour toute étude future une tâche d'articulation

purement laryngée (par exemple les coups de glotte). Cependant, la difficulté réside dans le fait d'impliquer certains muscles linguaux lors des mouvements laryngés, de par leurs connexions anatomiques avec la musculature laryngée extrinsèque.

- Par rapport à l'absence de phonémotopie trouvée dans notre étude, des analyses multivariées des patterns d'activité sont en cours dans le cadre d'une collaboration avec Michel Dojat et Mathieu Ruiz de l'Institut des Neurosciences de Grenoble. L'objectif de ce projet est de déterminer de possibles patterns d'activité propres aux différentes voyelles (ou traits phonétiques associés) perçues et produites, et en plus du cortex auditif dans d'autres régions d'intérêt, comme les régions primaires motrice, somatosensorielle, l'aire de Broca, le gyrus supramarginal. Il serait également possible de procéder à des analyses individuelles par régions d'intérêt dans ces mêmes régions, comme celles effectuées pour les régions du cortex sensorimoteur primaire lors des études 1.1 et 2.1.

Boucles correctives et codage des buts sensorimoteurs

Les études présentées dans le premier chapitre de cette thèse ont permis de préciser les réseaux neuronaux impliqués dans la réalisation de gestes articulatoires orofaciaux supralaryngés et lors de la production et de la perception de voyelles du Français.

En regard de ces résultats, une question importante concerne l'existence de régions sensibles à des mécanismes d'adaptation sensorimotrice. En effet, de nombreux modèles et études liées au contrôle moteur postulent l'existence de modèles internes dont le but serait de permettre l'apprentissage puis la réalisation correcte et fluide de mouvements et d'actions. Ces modèles internes coderaient les représentations et plan moteurs des actions ainsi que les conséquences sensorielles correspondantes, en d'autres termes le but de ces actions. La réalisation d'une action serait ainsi dépendante de ces modèles internes et de la mise en œuvre de boucles correctives en ligne, consistant en la comparaison des conséquences sensorielles prédites et réelles des actions. Dans ce cadre, certains modèles supposent l'existence de boucles prédictives et correctives en cascade, des niveaux supérieurs et inférieurs dans la hiérarchie des représentations d'actions en vue d'optimiser le contrôle des erreurs de prédictions (voir par exemple Kilner, 2007, pour plus de détails sur les processus de codage prédictif).

En ce qui concerne la production de la parole, plusieurs types de modèles neurobiologiques se basent sur des mécanismes parfois très similaires où les conséquences sensorielles prédites d'un acte de parole seraient évaluées par rapport au feedback sensoriel réel de manière à permettre, si nécessaire, des modifications rapides de production de parole (par exemple, Guenther, 2006; Price, Crinion et MacSweeney, 2011; Tian et Poeppel, 2010; Guenther et Vladusich, 2012; Hickok, 2012). Ce type de modèles neurobiologiques génératifs a également été proposé pour la perception de la parole où le système moteur permettrait par des mécanismes de simulation motrice et de copies d'efférence une désambiguïsation des informations phonétiques perçues au sein du cortex auditif (par exemple, Scott et Johnsrude, 2003; Callan et al., 2003; Wilson et Iacoboni, 2006; Scott et Wise, 2004; Skipper et al., 2007; Rauschecker et Scott, 2009; Hickok, Houde et Rong, 2011; Rauschecker, 2011).

Dans le cadre de possibles mécanismes adaptatifs reliés aux informations sensorielles et motrices perçues et/ou produites, nous avons utilisé un paradigme de répétition suppression lors des études 2.1 et 2.2 de manière à déterminer les régions possiblement impliquées dans le codage du but des actions perçues ou produites et dans des processus sensoriels prédictifs/adaptatifs.

Rappel des résultats

Adaptation neuronale aux buts moteurs orofaciaux (voir. Figure D.4). L'étude 2.1 consistait en la réalisation de mouvements orofaciaux simples (protrusion des lèvres, abaissement de la mandibule et rétraction de la langue) par séquences de 6 occurrences successives afin d'induire un effet de répétition suppression (décroissance du signal BOLD au fur et à mesure des répétitions). Les résultats montrent un effet de RS dans l'hémisphère gauche au niveau du gyrus supramarginal et du sulcus intrapariétal, du lobule pariétal supérieur (précuneus) et de la partie la plus dorsale du cortex prémoteur ventral.

Adaptation neuronale aux buts phonémiques produits et perçus (voir figure D.5). Dans l'étude 2.2, les participants percevaient et, dans la deuxième tâche, produisaient six mêmes voyelles de façon consécutive. Dans les deux tâches, les régions sensibles à l'effet de RS concernaient l'aire motrice supplémentaire et le cortex cingulaire antérieur, l'aire de Broca et le cortex prémoteur ventral, les lobules pariétaux supérieur et inférieur. Lors de la tâche de production, cet effet d'adaptation a également eu lieu dans le cervelet, le gyrus fusiforme et les aires extra-striées occipitales.

Adaptation sensorimotrice lors de la production des gestes orofaciaux et des voyelles (voir Figure D.4). La finalité des actions phonémiques et des actions orofaciales simples est évidemment de nature différente. Pour les actions phonémiques/vocaliques, le but est principalement communicatif. En vue d'une intelligibilité optimale, les paramètres et commandes contrôlés par le système moteur doivent donc être extrêmement précis. La réalisation des gestes orofaciaux n'implique évidemment pas un tel objectif et il est possible qu'un tel degré de précision des commandes motrices ne soit pas nécessaire.

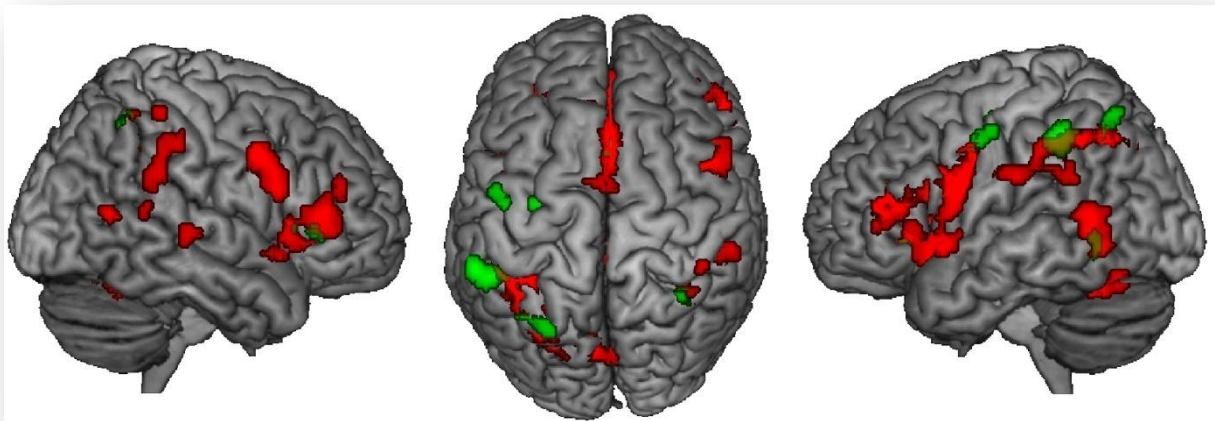


Figure D.4 : Superposition des régions sensibles à l'effet RS lors de la réalisation répétée des gestes articulatoires orofaciaux (en vert) et des voyelles (en rouge, seuil non corrigé, $p < .001$ et taille minimale des clusters de 25 voxels).

En ce qui concerne la tâche de production des voyelles, nos résultats sont en accord avec la littérature et les modèles génératifs cités plus avant qui proposent l'existence de boucles adaptatives/correctives sensorimotrices lors de la réalisation d'un acte de parole. En regard de nos résultats, les buts phonémiques des voyelles seraient ainsi représentés en termes sensorimoteurs : les représentations articulatoires seraient codées par le gyrus frontal inférieur et la partie ventrale du cortex prémoteur, les représentations auditives et somatosensorielles seraient elles codées par la partie postérieure du sulcus temporal supérieur et le lobule pariétal inférieur. Dans notre étude, comme chaque type de voyelle a été produite six fois à la suite, il est concevable qu'un modèle interne d'une voyelle donnée soit évoqué pour la première occurrence et mis en confrontation par le cervelet avec les retours sensoriels réels, mais qu'ensuite, étant donnée la grande facilité de la tâche, le système n'aurait pas besoin de surveiller une tâche aussi parfaitement maîtrisée. Les boucles correctives ne seraient alors que peu utiles et pourraient être désengagées afin de permettre une

allocation de ressources neuronales et attentionnelles vers d'autres régions cérébrales plus pertinentes pour la réalisation de la tâche (McKiernan et al., 2003)⁶¹.

En comparant les régions qui s'adaptent à la répétition des voyelles et des gestes orofaciaux, on observe un effet d'adaptation commun au niveau du cortex prémoteur ventral, du sulcus/lobule pariétal inférieur et du lobule pariétal supérieur. Ces deux régions ont été dans de précédentes études impliquées dans le codage du but d'actions manuelles (par exemple, Dinstein et al., 2007; Hamilton et Grafton, 2009; Kilner et al., 2009). Il a notamment été suggéré que ces régions soient impliquées dans des mécanismes de transformations visuo-motrices. Comme aucun retour visuel n'est possible dans le cas de gestes orofaciaux, l'apport de notre étude consiste ainsi à démontrer que ces régions seraient probablement impliquées dans des transformations autres que visuo-motrices, à savoir somatosensori-motrices. De manière importante, par rapport à la production des voyelles, les sites cérébraux sensibles à la répétition des gestes orofaciaux ne concernent ni le gyrus frontal inférieur ni les parties postérieures du gyrus et sulcus temporal supérieur. Il en ressort que ces régions ne codent pas le but des actions supralaryngées mais seraient donc propres au codage des représentations phonémiques. De plus, nos résultats montrent que le cervelet ne s'adapte pas pour les gestes orofaciaux simples, ce qui suggère une plus faible implication de cette structure en terme de coordination sensorimotrice pour ce type de mouvements.

Adaptation sensorimotrice lors de la production et la perception des voyelles (voir Figure D.5). Nos résultats montrent un effet d'adaptation phonémique commun aux tâches de perception et de production des voyelles dans des régions postérieures du gyrus/sulcus temporal supérieur. De manière intéressante, ce résultat pourrait ainsi indirectement appuyer l'hypothèse de Hickok et Poeppel par rapport aux débats concernant l'existence d'une voie auditive soit antérieure (Scott et Johnsrude, 2003; Rauschecker et Scott, 2009; Rauschecker, 2011) soit postérieure (Hickok et Poeppel, 2007) impliquée dans les processus de décodage acoustico-phonétique lors de la perception de la parole⁶². D'autre part, les résultats montrent aussi clairement des effets d'adaptation communs au niveau des aires prémotrices (gyrus frontal inférieur et cortex prémoteur ventral) et pariétales (lobule pariétal inférieur et sulcus intrapariétal).

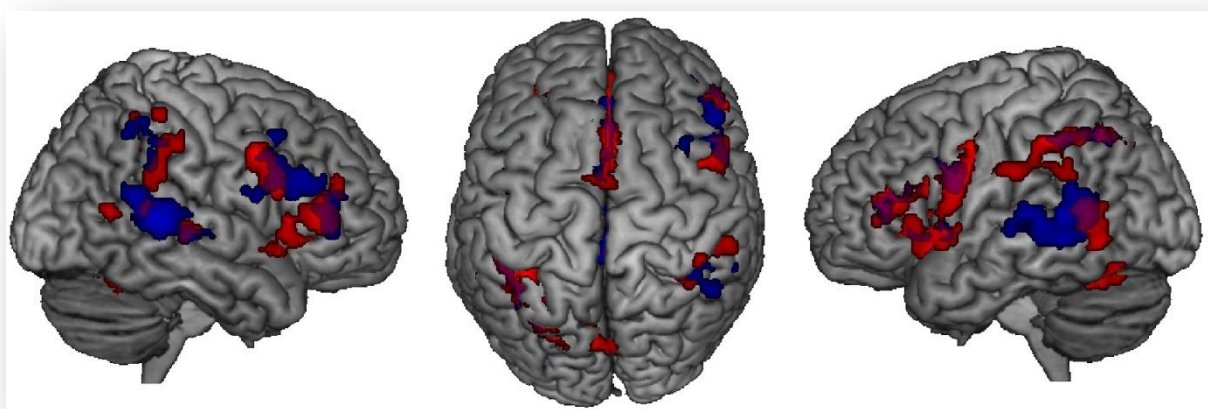


Figure D.5 : Superposition des régions sensibles à l'effet RS lors de l'écoute (en bleu) et de la production (en rouge) répétée des voyelles (seuil non corrigé, $p < .001$ et taille minimale des clusters de 25 voxels).

⁶¹ Il est intéressant de noter que les résultats de la tâche de production dans l'étude 1.2 sont cohérents avec cette interprétation. Nous observons en effet des déactivations de ces régions lors de la production simple de voyelles et sans induction d'adaptation mais seulement pour les voyelles fermées et mi-fermées (dont la réalisation est parfaitement maîtrisée). Par contre, ces régions sont sollicitées pour la production des voyelles mi-ouvertes, qui, dans la parole spontanée, ne sont pas naturellement produites dans un contexte isolé.

⁶² La possibilité d'une adaptation strictement acoustique semble écartée car les voyelles, bien qu'appartenant à une même catégorie phonémique, n'étaient pas identiques du point de vue acoustique.

Pris ensemble, nos résultats montrent ainsi des effets d'adaptation dans les deux tâches dans des régions sensorimotrices frontales, auditives temporales et somatosensorielles pariétales. Cette étude suggère ainsi que le but des actions phonémiques perçues et produites est bien de nature sensorimotrice.

Boucles régulatrices et cartes d'erreurs. Il ressort de nos résultats une double dissociation fonctionnelle entre l'implication des lobules pariétaux inférieur/supérieur et la jonction temporo-pariétale. Les régions pariétales sont sensibles à l'effet de RS dans les trois tâches et notamment lors de la réalisation des gestes orofaciaux simples (nécessitant un rétrocontrôle de nature somatosensorielle mais pas auditif) tandis que la jonction temporo-pariétale ne montre un effet de RS que lors de la production et de la perception des voyelles (ce qui suggère fortement un rôle de cette région dans le codage auditif phonémique).

De manière importante, pour ces régions, les patterns d'adaptation entre les six occurrences diffèrent cependant dans les trois tâches, ce qui souligne la nature active des mécanismes adaptatifs régulateurs. Notre interprétation, loin d'être exhaustive, est la suivante. Dans la tâche de production de voyelles, la première occurrence nécessite une mise en place d'un programme moteur approprié et bien précis, connu et maîtrisé (modèle inverse). Une fois ce plan moteur établi et les corrections sensorimotrices (cartes d'erreurs) effectuées lors de la première production, il n'y a plus nécessité de répéter ces processus (d'autant plus que les participants avaient connaissance du paradigme de répétition des voyelles ou des gestes). Cela pourrait expliquer la chute brusque du signal BOLD de la première occurrence à la seconde et à la stabilité du signal de la seconde à la dernière occurrence dans les diverses régions frontales, temporales et pariétales ainsi que dans le cervelet. La décroissance linéaire d'activité dans les mêmes régions frontales, pariétales et temporales lors de la tâche de perception passive suggère également que les boucles correctives sensorimotrices se désimpliquent au fur et à mesure de la répétition de la même voyelle. Par comparaison avec la décroissance brutale de la première à la seconde occurrence observée lors de la tâche de production, cette décroissance linéaire pourrait en fait impliquer de manière conjointe un effet de RS et des effets de baisse attentionnelle (l'augmentation d'activité observée lors de la dernière occurrence appuie d'ailleurs cette idée). Ce possible effet combiné de RS et de baisse attentionnelle pourrait également expliquer la décroissance linéaire observée dans le cortex prémoteur ventral et les régions pariétales postérieures lors de la réalisation de gestes orofaciaux. En effet, comparée à la production des voyelles, il est possible qu'un degré moindre de précision motrice soit nécessaire à la réalisation des gestes orofaciaux. Cela pourrait ainsi se traduire par un désengagement moins brutal des régions motrices et pariétales impliquées dans les processus de régulation.

Limites et perspectives

Boucles correctives et buts des actions orofaciales

- Les effets de RS observés dans les régions prémotrices et pariétales postérieures dans l'étude 2.1 suggèrent que cet effet ne dépend pas de la réplication précise et identique d'une action d'un essai à l'autre mais bien du codage du but de l'action. Comme discuté précédemment, il aurait été néanmoins important de contrôler les paramètres cinématiques et musculaires lors de la réalisation des gestes supralaryngés. Face à une conception hiérarchique du contrôle moteur, une perspective majeure à cette étude serait d'induire des variations des mouvements réalisés (par exemple, selon la dimension antérieur-postérieur, la dimension gauche-droite, ou en fonction de la rapidité du mouvement) et d'ainsi pouvoir tester les régions sensibles à l'effet de RS en fonction de ces paramètres cinématiques, musculaires et du but de l'action orofaciale.
- Une autre question méthodologique concerne le nombre limité des essais par tâche motrice. Pour cette raison, l'effet de RS a été modélisé indépendamment de l'effecteur. Bien que nos résultats suggèrent que le codage du but des actions orofaciales soit effectué dans les aires prémotrices et pariétales gauches, il n'est pas à exclure que l'effet RS observé soit relié

principalement à l'une des tâches motrices. Des études complémentaires sont donc nécessaires afin de déterminer plus spécifiquement l'effet de RS pour chacun des mouvements linguaux, labiaux ou mandibulaires.

- Afin de préciser les liens entre représentations des buts sensorimoteurs et les boucles correctives, on pourrait ajouter un facteur de perturbation motrice (voir par exemple l'étude de Golfinopoulos et al., 2011) afin de voir la concordance entre but moteur et correction de cette perturbation motrice.
- Dans le cas de l'abaissement de la mandibule ou de la protrusion des lèvres, une étude similaire consistant en l'observation de ces gestes pourrait permettre de démontrer un couplage fonctionnel entre systèmes de perception et de production des gestes orofaciaux.
- Enfin, comme présenté précédemment, des analyses ou études supplémentaires pourraient concerner la latéralisation, la connectivité fonctionnelle entre régions et le décours temporel des régions sensibles à l'effet de RS.

Boucles correctives et buts phonémiques des voyelles perçues et produites

- Une limite majeure de l'étude 2.2 est le nombre insuffisant d'essais qui empêche la comparaison de l'effet d'adaptation par voyelle et même par trait phonétique. Pour vérifier cela, il faudrait limiter le nombre de voyelles utilisées ou le nombre de répétitions afin de tester plus précisément de possibles différences entre traits phonétiques ou entre types de voyelles.
- Dans le cadre des modèles internes, il serait également intéressant d'étudier plus finement un possible apprentissage des relations sensorimotrices pour les voyelles mi-ouvertes, moins maîtrisées, ou de manière plus générale pour des unités de parole non-apprises.
- Dans la présente étude, la tâche de perception des voyelles étaient réalisée avant celle de production de manière à limiter de possibles effets de récapitulation subvocale. Néanmoins, par rapport à la critique de Lingnau, Gesierich et Caramazza (2008), selon qui l'exploration d'un système de neurones miroirs chez humain nécessite une démonstration d'un effet d'adaptation cross-modal dans les deux sens action-perception et perception-action, il pourrait être intéressant de mêler écoute et production des voyelles lors d'une même session entre deux essais successifs. Il serait de plus possible de croiser les facteurs tâches et voyelles/traits entre paires d'essais (même tâche et même trait, même tâche mais traits différents, etc). Cette proposition permettrait de déterminer plus finement les populations neuronales adaptées.
- Comme cité plus avant, une étude sur les connectivités fonctionnelles et le décours temporel des régions sensibles à l'effet de RS pourrait apporter de précieuses informations quant au réseau neural gouvernant les boucles correctives et les fonctions sous-jacentes (liens entre cervelet et régions sensorimotrices, cartes d'erreurs, mises à jour des plans moteurs).
- Enfin, en regard des modèles neurobiologiques cités précédemment, des études impliquant dans une même expérience effet d'adaptation et/ou perturbation sensorimotrice (auditive, motrice ou somatosensorielle) permettrait de préciser le rôle de ces régions au sein de ces réseaux.

Implication causale de la voie dorsale lors de la perception de la parole

Si les résultats observés lors des études IRMf présentées dans les deux premières parties de la présente thèse démontrent l'implication de la voie dorsale lors de la perception et de la production des voyelles du Français, ces résultats sont par nature corrélationnels et ne permettent pas de démontrer un rôle médiateur, fonctionnel, des régions activées.

Nous discuterons dans cette dernière sous-partie des résultats des études présentées dans la troisième partie de cette thèse qui avaient pour but de déterminer l'implication fonctionnelle et causale des régions sensorielles et motrices de la voie dorsale lors de la perception d'unités de

parole, par l'utilisation de la technique de stimulation magnétique transcrânienne et de paradigmes d'apprentissage/adaptation moteur et auditif à court-terme.

Rappel des résultats

Lors de l'étude 3.1, nous avons utilisé la technique de stimulation magnétique transcrânienne et induit un effet d'adaptation auditive afin de vérifier l'implication possible des régions de la voie dorsale dans la catégorisation de sons de parole ambigus. Nos résultats suggèrent un rôle causal du cortex prémoteur ventral, de la partie dorsale du gyrus supramarginal et de la partie postérieure du gyrus supérieur temporal dans la perception catégorielle de la parole.

De manière plus spécifique, après stimulation, nous avons observé dans la seconde partie de la tâche (BIN2) le maintien de l'effet d'adaptation sélective alors que celui-disparaissait lors d'une condition contrôle sans stimulation réelle. Les différences observées lors de la seconde partie de la tâche peuvent selon nous s'expliquer par un effet progressif de la stimulation, fonction du nombre de pulses. L'excitabilité induite par stimulation des populations neuronales permettrait ainsi le maintien de l'effet d'adaptation sélective. De manière importante, aucune différence n'a été observée en fonction de la région stimulée. Cela suggère que ces trois régions fonctionnent de concert au sein de la voie dorsale, bien qu'une stimulation d'une région contrôle non potentiellement impliquée aurait permis de confirmer cette hypothèse. Même si certains résultats restent à élucider, cela n'amoindrit pas les résultats qui, en accord avec de précédentes études, montrent une implication de la voie dorsale dans la catégorisation des sons de parole et sa sensibilité à l'effet d'adaptation auditive.

De manière parallèle, nous avons utilisé lors de l'étude 3.2 un entraînement moteur labial ou lingual de manière à tester si une modulation des activités motrices induite par cet entraînement pouvait modifier la catégorisation de syllabes présentées auditivement. Les biais de réponse observés, reliés spécifiquement à l'articulateur utilisé lors de la tâche motrice et à la syllabe perçue, suggèrent que le système moteur jouerait un rôle causal spécifique dans les processus de catégorisation lors de la perception de la parole (voir d'Ausilio et al., 2009 pour une conclusion similaire).

Prises ensemble, ces deux études contribuent ainsi à mettre en valeur le rôle fonctionnel du système moteur et de la voie dorsale dans la perception de la parole.

Limites et perspectives

Les limites et perspectives de ces études concernent principalement certains aspects méthodologiques utilisés lors de ces études :

- Lors de l'étude 3.1, le fait d'alterner les différents effets d'adaptation auditive et de perturbation corticale des différentes régions se base sur la présupposition que l'induction de ces deux types d'effets est temporellement circonscrite. Il serait néanmoins important de tester réellement la durée de ces effets afin de contrôler plus finement leurs conséquences et d'amoindrir le risque d'after-effets qui peuvent perdurer entre conditions.
- Un paramètre à étudier prochainement serait le décours temporel d'activation de la voie dorsale lors de la perception des sons de parole. Pour ce faire, il s'agirait de faire écouter des syllabes ambiguës, de manière similaire à l'étude 3.1, mais de stimuler les mêmes sites de la voie dorsale selon plusieurs latences temporelles.
- De plus, un paradigme similaire permettrait d'étudier à la fois la tâche de perception (de syllabes ambiguës) et de production (avec modification du feedback auditif). D'autre part, dans le cadre d'une implication de la voie dorsale dans les mécanismes d'imitation volontaire et de convergence phonétique, une autre étude pourrait consister en une tâche d'imitation volontaire ou automatique de la parole.

- Enfin, un projet postdoctoral récemment soumis vise à étudier les sites d'intégration sensorimotrice pour deux types d'actions vocales, parlée et chantée, aussi bien chez des chanteurs que chez des participants non experts, et en présence d'une modification du feedback auditif ou d'une perturbation motrice. Ce projet permettrait de comparer les mécanismes de contrôle de deux systèmes sensorimoteurs supposés similaires du point de vue articulaire mais impliquant probablement des boucles correctives et de régulation différentes.

Conclusion

L'objectif de cette thèse était d'interroger la possible nature sensorimotrice des représentations cérébrales des unités de parole. Bien que nous n'ayons pu mettre en évidence de patterns phonémotopiques associés aux voyelles du Français, nos études démontrent l'existence d'un couplage fonctionnel des systèmes de perception et de production de la parole et appuient ainsi l'hypothèse d'un codage sensorimoteur distribué des représentations de la parole, même dans le cas de la perception d'unités élémentaires ne nécessitant à priori qu'une implication faible du système moteur dans les processus de décodage et désambiguïsation phonétique.

Nos travaux ont en effet permis de mettre en évidence l'implication de régions auditives, motrices et somatosensorielles aussi bien lors de la production que de la perception des voyelles. En témoignent les activations temporales, frontales et pariétales lors des tâches de perception et de production des voyelles, ainsi que la modulation d'activité de certaines de ces régions pour les voyelles mi-ouvertes, moins maîtrisées en contexte isolé. La mise en évidence d'un effet d'adaptation pour des régions motrices, auditives et somatosensorielles lors de l'écoute ou de la réalisation répétée d'une même voyelle suggère de plus l'existence de boucles sensorimotrices communes, impliquées dans des mécanismes de régulation et de contrôle en ligne des voyelles perçues et produites. Enfin, nous avons pu démontrer le rôle causal et fonctionnel des régions sensorielles et motrices de la voie dorsale lors de la perception de sons de parole.

Pris ensemble, les études présentées dans cette thèse suggèrent fortement un couplage entre système de perception et de production et, par cela, supportent la proposition d'une nature intrinsèquement sensorimotrice des unités de parole telle que défendue dans la Théorie de la Perception pour le Contrôle de l'Action élaborée par les chercheurs de notre laboratoire.

- Bibliographie -

Bibliographie

- Amunts K, Jäncke L, Mohlberg H, Steinmetz H & Zilles K (2000). Interhemispheric asymmetry of the human motor cortex related to handedness and gender. *Neuropsychologia*, 38(3):304-12.
- Arbib MA (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28(2):105-124.
- Atal BS, Chang JJ, Mathews MV & Tukey JW (1978). Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique, *Journal of the Acoustical Society of America*, 63, 1535-1555.
- Aubanel V (2011). Variation *phonologique régionale en interaction conversationnelle*. Thèse de Doctorat, Aix-Marseille Université.
- Aubanel V & Nguyen N (2010). Automatic recognition of regional phonological variation in conversational interaction. *Speech Communication*, 52:577-586.
- Baciu, M (2011). *Introduction à la Neuroanatomie Fonctionnelle*, De Boeck Université. Louvain la Neuve, Belgique.
- Baddeley A (1986). *Working Memory*. Oxford, Clarendon Press
- Basirat A, Sato M, Schwartz JL, Kahane P & Lachaux JP (2008). Parieto-frontal gamma band activity during the perceptual emergence of speech forms. *NeuroImage*, 42(1):404-13.
- Basirat A, Schwartz JL & Sato M (2012). Perceptuo-motor interactions in the perceptual organization of speech: evidence from the verbal transformation effect. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 367(1591):965-976.
- Basso A, Casati G & Vignolo LA (1977). Phonemic identification defects in aphasia. *Cortex*, 13:84-95.
- Bear MF, Connors BW & Paradiso MA (2002). *Neurosciences : à la découverte du cerveau*. Editions Pradel.
- Belin P & Zatorre RJ (2000). 'What', 'where' and 'how' in auditory cortex. *Nature Neuroscience*, 3(10):965-6.
- Besle J, Caclin A, Mayet R, Delpuech C, Lecaigard F, Giard MH & Morlet D (2007). Audiovisual events in sensory memory. *Journal of Psychophysiology*, 21(3-4), 231-238.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao, SM & Cox RW (1999). Conceptual processing during the conscious resting state: A functional MRI study. *Journal of Cognitive Neuroscience*, 11, 80-93.
- Binder JR, Liebenthal E, Possing ET, Medler DA & Ward BD (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*, 7:295-301.
- Birn RM, Bandettini PA, Cox RW & Shaker R (1999). Event-related fMRI of tasks involving brief motion. *Human Brain Mapping*, 7(2):106-114.
- Blakemore SJ & Frith C (2003). Self-awareness and action. *Current opinion in neurobiology*, 13(2):219-24.
- Blakemore SJ, Wolpert DM & Frith CD (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7):635-40.
- Blumstein SE & Stevens KN (1979). Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, 66, 1001-1017.
- Blumstein SE, Cooper WE, Zurif EB & Caramazza A (1977). The perception and production of voice

- onset time in aphasia. *Neuropsychologia*, 15:371-383.
- Boatman DF & Miglioretti DL (2005). Cortical sites critical for speech discrimination in normal and impaired listeners. *Journal of Neuroscience*, 25(23):5475–5480.
- Boatmann DF (2004). Cortical bases of speech perception: evidence from functional lesion studies. *Cognition*, 92:47-65.
- Boemio A, Fromm S, Braun A & Poeppel D (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, 8, 389–395.
- Bohland JW & Guenther FH (2006). An fMRI investigation of syllable sequence production. *NeuroImage*. 32(2):821-41.
- Bonilha L & Fridriksson J (2009). Subcortical damage and white matter disconnection associated with non-fluent speech. *Brain*, 132(Pt 6):e108.
- Bonilha L, Moser D, Rorden C, Baylis GC & Fridriksson J (2006). Speech apraxia without oral apraxia: can normal brain function explain the physiopathology? *Neuroreport*, 17(10):1027-31.
- Bonini L, Rozzi S, Serventi FU, Simone L, Ferrari PF & Fogassi L (2010). Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. *Cerebral Cortex*, 20(6):1372-85.
- Bonini L, Serventi FU, Simone L, Rozzi S, Ferrari PF & Fogassi L (2011). Grasping neurons of monkey parietal and premotor cortices encode action goals at distinct levels of abstraction during complex action sequences. *Journal of Neuroscience*, 31(15):5876-86.
- Booth JR, Cho S, Burman DD & Bitan T (2007). Neural correlates of mapping from phonology to orthography in children performing an auditory spelling task. *Developmental science*, 10(4):441-51.
- Boudreau S, Romaniello A, Wang K, Svensson P, Sessle BJ & Arendt-Nielsen L (2007). The effects of intra-oral pain on motor cortex neuroplasticity associated with short-term novel tongue-protrusion training in humans. *Pain*, 132(1-2):169-78.
- Bridgeman B (2007). Efference copy and its limitations. *Computers in Biology and Medicine*. 37(7):924-9.
- Britton B, Blumstein SE, Myers EB & Grindrod C (2009). The role of spectral and durational properties on hemispheric asymmetries in vowel perception. *Neuropsychologia*, 47(4):1096-106.
- Brodmann K (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*, Johann Ambrosius Barth Verlag, Leipzig, Germany.
- Brooks DJ (1995). The role of the basal ganglia in motor control: contributions from PET. *The Journal of Neuroscience*. 128:1.
- Browman C & Goldstein L (1989). Articulatory gestures as phonological units. *Phonology*, 6, 201-251.
- Brown S, Ingham RJ, Ingham JC, Laird AR & Fox PT (2005). Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*, 25(1):105-17.
- Brown S, Ngan E & Liotti M (2008). A larynx area in the human motor cortex. *Cerebral Cortex*, 18(4):837- 845.
- Brown S, Laird AR, Pfordresher PQ, Thelen SM, Turkeltaub P & Liotti M (2009). The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain and cognition*, 70(1):31-41.
- Brunswik, E (1956). *Perception and the representative design of psychological experiments* (2nd ed., rev. & enl.), Berkeley: University of California Press.
- Burton MW, Small SL & Blumstein SE (2000). The Role of Segmentation in Phonological Processing: An fMRI Investigation. *Journal of Cognitive Neuroscience*, 12(4):679-90.
- Callan D, Callan A, Gamez M, Sato MA & Kawato M (2010). Premotor cortex mediates perceptual performance. *NeuroImage*, 51(2):844-58.
- Callan DE, Jones JA, Callan AM & Akahane-Yamada R (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic

- phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage*, 22:1182-1194.
- Callan DE, Jones JA, Munhall KG, Callan AM, Kroos C & Vatikiotis-Bateson E (2003). Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport*, 14:2213-2217.
- Calvert GA & Campbell R (2003). Reading speech from still and moving faces: The neural substrates of visible speech. *Journal of Cognitive Neuroscience*, 15(1):57-70.
- Caplan D, Gow D & Makris N (1995). Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology*, 45:293-298.
- Carota F, Posada A, Harquel S, Delpuech C, Bertrand O & Sirigu A (2010). Neural dynamics of the intention to speak. *Cerebral Cortex*, 20(8):1891-7.
- Cattaneo L, Fabbri-Destro M, Boria S, Pieraccini C, Monti A, Cossu G. & Rizzolatti, G (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proceedings of the National Academy of Sciences of the United States of America*, 104(45):17825-30.
- Cattaneo L (2010). Tuning of ventral premotor cortex neurons to distinct observed grasp types: a TMS-priming study. *Experimental Brain Research*, 207(3-4):165-172.
- Cattaneo L, Barchiesi G, Tabarelli D, Arfeller C, Sato M & Glenberg AM (2011). One's motor performance predictably modulates the understanding of others' actions through adaptation of premotor visuo-motor neurons. *Social Cognitive and Affective Neuroscience*, 6(3):301-310.
- Cattaneo L, Sandrini M & Schwarzbach J (2010). State-dependent TMS reveals a hierarchical representation of observed acts in the temporal, parietal, and premotor cortices. *Cerebral Cortex*, 20(9):2252-2258.
- Cattaneo Z & Silvanto J (2008). Investigating visual motion perception using the transcranial magnetic stimulation-adaptation paradigm. *Neuroreport*, 19:1423-1427.
- Cattaneo Z, Devlin JT, Salivini F, Vecchi T & Silvanto J (2010). The causal role of category-specific neuronal representations in the left ventral premotor cortex (PMv) in semantic processing. *NeuroImage*, 49(3):2728-2734.
- Chang SE, Kenney MK, Loucks TM, Poletto CJ & Ludlow CL (2009). Common neural substrates support speech and non-speech vocal tract gestures. *NeuroImage*. 47(1):314-25.
- Chevillet M, Riesenhuber M & Rauschecker JP (2007). Functional localization of the auditory "what" stream hierarchy. *Society for Neuroscience abstracts* 33, 174.9.
- Chong TT, Cunnington R, Williams MA, Kanwisher N & Mattingley JB (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current Biology*. 18(20):1576-80.
- Christoffels IK, Formisano E & Schiller NO (2007). Neural correlates of verbal feedback processing: an fMRI study employing overt speech. *Human Brain Mapping* 28(9):868-879.
- Christoffels IK, van de Ven V, Waldorp LJ, Formisano E & Schiller NO (2011). The sensory consequences of speaking: parametric neural cancellation during speech in auditory cortex. *PLoS One*. 6(5):e18307
- Cooper WE & Lauritsen MS (1974). Feature processing in the perception and production of speech. *Nature*, 252, 121-123.
- Cooper WE (1979). *Speech perception and production: Studies in selective adaptation*. Norwood, NJ: Ablex.
- Corfield DR, Murphy K, Josephs O, Fink GR, Frackowiak RS, Guz A, Adams L & Turner R. (1999). Cortical and subcortical control of tongue movement in humans: a functional neuroimaging study using fMRI. *Journal of Applied Physiology*. 86(5):1468-77.
- Cowie RI & Douglas-Cowie E. (1983). Speech production in profound post-lingual deafness. In *Hearing science and hearing disorders* (ME Lutman & MP Haggard, editors), pp. 183-231. New York: Academic Press.

- Cowie RI & Douglas-Cowie E. (1992). *Postlingually acquired deafness: speech deterioration and the wider consequences*. New York: Mouton de Gruyter.
- Critchley HD, Corfield DR, Chandler MP, Mathias CJ & Dolan RJ (2000). Cerebral correlates of autonomic cardiovascular arousal, a functional neuroimaging investigation in humans. *The Journal of Physiology*, 523 Pt 1:259-70.
- Curio G, Neuloh G, Numminen J, Jousmaki V & Hari R (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*, 9 :183–191.
- d'Ausilio A, Bufalari I, Salmas P & Fadiga L (2012). The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*, 48(7):882-7.
- d'Ausilio A, Jarmolowska J, Busan P, Bufalari I & Craighero L (2011). Tongue corticospinal modulation during attended verbal stimuli: priming and coarticulation effects. *Neuropsychologia*, 49(13):3670-6
- d'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C & Fadiga L (2009). The motor somatotopy of speech perception. *Current Biology*, 19(5):381-385.
- Decety J & Grèzes J (2006). The power of simulation: imagining one's own and other's behavior. *Brain Research*, 1079(1):4-14.
- Decety J & Sommerville JA (2003). Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences*, 7(12):527-33.
- Dehaene S (2008). Cerebral constraints in reading and arithmetic: Education as a «neuronal recycling» process. In *The educated brain*, pages pp 232-248. Cambridge University Press.
- Dehaene S, Naccache L, Cohen L, Bihan DL, Mangin JF, Poline JB & Rivière D (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4(7):752-758.
- Dehaene S, Pegado F, Braga LW, Ventura P, Nunes Filho G, Jobert A, Dehaene-Lambertz G, Kolinsky R, Morais J & Cohen L (2010). How learning to read changes the cortical networks for vision and language. *Science*, 330(6009):1359-64.
- Delvaux V & Soquet A (2007). The Influence of Ambient Speech on Adult Speech Productions through Unintentional Imitation. *Phonetica*, 64, 145-173.
- Démonet JF, Price C, Wise R & Frackowiak RS (1994). A PET study of cognitive strategies in normal subjects during language tasks. Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain*, 117 (Pt 4):671-82.
- Démonet JF, Chollet F, Ramsay S, Cardebat D, Nespoulous JL, Wise R, Rascol A & Frackowiak R (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115, 1753-1768.
- Démonet JF, Thierry G & Cardebat D (2005). Renewal of the neurophysiology of language: Functional neuroimaging. *Physiological Reviews*, 85, 49-95.
- Desmurget M, Reilly KT, Richard N, Szathmari A, Mottolese C & Sirigu A (2009). Movement intention after parietal cortex stimulation in humans. *Science*, 324(5928):811-3.
- Desmurget M & Sirigu A (2009). A parietal-premotor network for movement intention and motor awareness. *Trends in Cognitive Science*, 13,10:411-419.
- Devinsky O, Morrell MJ & Vogt BA (1995). Contributions of anterior cingulate cortex to behaviour. *Brain*, 118 (Pt 1):279-306.
- Dhanjal NS, Handunnetthi L, Patel MC & Wise RJ (2008). Perceptual systems controlling speech production. *Journal of Neuroscience*, 28(40):9969-75.
- Di Pellegrino G, Fadiga L, Fogassi L, Gallese V & Rizzolatti G (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research* 91:176-180.
- Diehl RL & Kluender KR (1989). On the objects of speech perception. *Ecological Psychology*, 2:121–44
- Diehl RL, Lotto AJ & Holt LL (2004). *Speech Perception*. Annual Review of Psychology, 55:149-179.
- Dinstein I, Hasson U, Rubin N & Heeger DJ (2007). Brain areas selective for both observed and

- executed movements. *Journal of Neurophysiology*. 98(3):1415-27.
- Dooling RJ, Best CT & Brown SD (1995). Discrimination of synthetic full-formant and sinewave /ra-la/ continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*). *Journal of the Acoustical Society of America*, 97:1839-46
- Doya K (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex. *Neural Networks*, 12(7-8):961-974.
- Dresel C, Castrop F, Haslinger B, Wohlschlaeger AM, Hennenlotter A & Ceballos-Baumann AO (2005). The functional neuroanatomy of coordinated orofacial movements: sparse sampling fMRI of whistling. *NeuroImage*, 28(3):588-97.
- Dronkers NF, Plaisant O, Iba-Zizen MT & Cabanis EA (2007). Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130(Pt 5):1432-41.
- Dronkers NF (1996). A new brain region for coordinating speech articulation. *Nature*, 384(6605):159-61.
- Durand J, Laks B & Lyche C (2002). La phonologie du français contemporain: usages, variétés et structure. In C. Pusch & W. Raible (eds.), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics -Corpora and Spoken Language*. Tübingen: Gunter Narr Verlag, pp. 93-106.
- Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K & Zilles K (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25:1325-1335.
- Eimas PD & Corbit JD (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 4:99-109.
- Epstein, Melissa, Narineh Hacopian & Peter Ladefoged (eds.) (2002). *Dissection of the Speech Production Mechanism*. *UCLA Working Papers in Phonetics*, 102.
- Estow S, Jamieson J & Yates J (2007). Self-monitoring and mimicry of positive and negative social behaviors. *Journal of Research in Personality*, 41:425-433.
- Fabbro F, Clarici A & Bava A (1996). Effects of left basal ganglia lesions on language production. *Perceptual and motor skills*. 82(3 Pt 2):1291-8.
- Fadiga L, Craighero L & Olivier E (2005). Human motor cortex excitability during the perception of others' action. *Current Opinion in Neurobiology*, 15:213-18,
- Fadiga L, Craighero L, Buccino G & Rizzolatti G (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, 15:399-402.
- Ferrari PF, Gallese V, Rizzolatti G & Fogassi L (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*, 17:1703-1714.
- Fodor JA (1983). *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass.: MIT Press
- Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F & Rizzolatti, G (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308:662-667.
- Formisano E, De Martino F, Bonte M & Goebel R (2008). "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science*, 322:970-973.
- Fort M (2011). *L'accès au lexique dans la perception audiovisuelle et visuelle de la parole*. Thèse de Doctorat, Grenoble Université.
- Fowler CA & Rosenblum LD (1990). "Duplex perception: A comparison of monosyllables and slamming doors". *Journal of experimental psychology. Human perception and performance* 16 (4):742-754.
- Fowler CA (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3-28.
- Fowler CA (1996). Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America*

- Francis BA & Wonham WM (1976). The internal model principle of control theory. *Automatica*, 12, 457-651
- Friston K (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 360(1456):815-36.
- Friston K (2011). What is optimal about motor control? *Neuron*. 72(3):488-98.
- Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS & Turner R (1995). Analysis of fMRI time-series revisited. *NeuroImage* 2(1):45-53.
- Friston KJ, Holmes AP, Price CJ, Buchel & C Worsley KJ (1999). Multisubject fMRI studies and conjunction analyses. *NeuroImage*, 10:385-396.
- Friston KJ, Holmes, AP & Worsley KJ (1999). How many subjects constitute a study? *NeuroImage*, 10:1-5.
- Fu CH, Vythelingum GN, Brammer MJ, Williams SC, Amaro E Jr, Andrew CM, Yágüez L, van Haren NE, Matsumoto K & McGuire PK (2006). An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal feedback. *Cerebral Cortex*, 16(7), 969-977.
- Galantucci B, Fowler CA & Turvey MT (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13(3):361-377.
- Gallese V, Fadiga L, Fogassi L & Rizzolatti G (1996). Action recognition in the premotor cortex. *Brain* 119:593-609.
- Gallese V, Murata A, Kaseda M, Niki N & Sakata H (1994). Deficit of hand preshaping after muscimol injection in monkey parietal cortex. *Neuroreport*, 5(12):1525-9.
- Ganesh G, Haruno M, Kawato M & Burdet E (2010). Motor memory and local minimization of error and effort, not global optimization, determine motor behavior. *Journal of Neurophysiology*, 104(1):382-90.
- Gardner E, Ro J, Babu K & Ghosh S (2007). Neurophysiology of prehension. II. Response diversity in primary somatosensory (S-I) and motor (M-I) cortices. *Journal of Neurophysiology*, 97(2):1656-70.
- Genovese CR, Lazar NA & Nichols T (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870-8.
- Gentilucci M & Bernardis P (2007). Imitation during phoneme production. *Neuropsychologia*, 45(3):608-15.
- Gentilucci M & Cattaneo L (2005). Automatic audiovisual integration in speech perception. *Experimental Brain Research*, 167(1):66-75.
- Gentilucci M, Fogassi L, Luppino G, Matelli M, Camarda R & Rizzolatti G (1988). Functional organization of inferior area 6 in the macaque monkey. I. Somatotopy and the control of proximal movements. *Experimental Brain Research*, 71(3):475-90.
- Gentilucci M & Corballis MC (2006). From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews*, 30:949-960.
- Georgopoulos AP, Schwartz AB & Kettner RE (1986). Neuronal population coding of movement direction. *Science*, 233(4771):1416-9.
- Geyer S, Ledberg A, Schleicher A, Kinomura S, Schormann T, Bürgel U, Klingberg T, Larsson J, Zilles K & Roland PE (1996). Two different areas within the primary motor cortex of man. *Nature*, 382(6594):805-7.
- Geyer S, Schleicher A & Zilles K (1999). Areas 3a, 3b, and 1 of human primary somatosensory cortex: 1. Microstructural organization and interindividual variability. *NeuroImage*, 10:63– 83.
- Ghosh SS, Tourville JA & Guenther FH (2008). A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *Journal of Speech, Language, and Hearing Research*, 51(5):1183-202.
- Gibson JJ (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin

- Gibson JJ (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin
- Giles H, Coupland N & Coupland J (1991). Accommodation theory: Communication, context, and consequence. In *Contexts of Accommodation: Developments in Applied Sociolinguistics* (editors: Giles H, Coupland N & Coupland J editors), 1-68. Cambridge University Press, Cambridge, UK
- Glenberg AM, Lopez-Mobilia G, McBeath M, Toma M, Sato M & Cattaneo L (2010). Knowing beans: Human mirror mechanisms revealed through motor adaptation. *Frontiers in Human Neuroscience*, 4:206.
- Golfinopoulos E, Tourville JA, Bohland JW, Ghosh SS, Nieto-Castanon A & Guenther FH (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *NeuroImage*, 55(3):1324-38.
- Goodale MA & Milner AD (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1):20-5.
- Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M (2012). Functional MRI assessment of orofacial articulators: neural correlates of lip, jaw, larynx and tongue movements. *Human Brain Mapping*, 33(10):2306-21.
- Gracco VL & Abbs JH (1985). Dynamic control of the perioral system during speech: kinematic analyses of autogenic and nonautogenic sensorimotor processes. *Journal of Neurophysiology*, 54(2):418-32.
- Gracco VL & Tremblay P, Pike GB (2005). Imaging speech production using fMRI. *NeuroImage*, 26:294-301.
- Gracco, VL & Löfqvist A (1994). Speech motor organization and control: Evidence from lip, jaw, and laryngeal interactions. *Journal of Neuroscience*, 14, 6585-6597.
- Grafton ST, Fagg AH & Arbib MA (1998). Dorsal premotor cortex and conditional movement selection: A PET functional mapping study. *Journal of Neurophysiology*, 79(2):1092-7.
- Grafton ST & Hamilton AF (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4):590-616.
- Grill-Spector K, Henson R & Martin A (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10:14–23.
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y & Malach R (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, 24(1):187-203.
- Grill-Spector K & Malach R (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, 107(1-3):293-321.
- Guenther FH, Ghosh SS & Tourville JA. 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96(3):280-301.
- Guenther FH, Hampson M & Johnson D (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, 105(4):611-33.
- Guenther FH & Vladusich T (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, 25(5):408–422
- Guenther FH (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, 39:350-365.
- Gusnard DA, Akbudak E, Shulman GL & Raichle ME (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*. 98(7):4259-64.
- Haaland KY, Harrington DL & Knight RT (2000). Neural representations of skilled movement. *Brain*, 123, 2306–2313.
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM & Bowtell RW (1999). "Sparse" temporal sampling in auditory fMRI. *Human Brain Mapping*, 7(3):213-23.

- Hamdy S, Mikulis DJ, Crawley A, Xue S, Lau H, Henry S & Diamant NE (1999). Cortical activation during human volitional swallowing: an event-related fMRI study. *The American Journal of Physiology*, 277(1):G219-25.
- Hamilton AF & Grafton ST (2009). Repetition suppression for performed hand gestures revealed by fMRI. *Human Brain Mapping*, 30(9):2898-906.
- Hamilton AF & Grafton ST (2006). Goal representation in human anterior intraparietal sulcus. *Journal of Neuroscience*, 26(4):1133-7.
- Hartwigsen G, Baumgaertner A, Price CJ, Koehnke M, Ulmer S & Siebner HR (2010). Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the National Academy of Sciences of the United States of America*, 107(38):16494-9.
- Hashimoto Y & Sakai KL (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study. *Human Brain Mapping*, 20:22-28.
- Hasson U, Skipper JI, Nusbaum HC & Small SL (2007). Abstract coding of audiovisual speech: beyond sensory representation. *Neuron*, 56(6):1116-26.
- Heinks-Maldonado TH, Nagarajan SS & Houde JF (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *NeuroReport*, 17(13):1375-1379.
- Helmholtz H (1866/1962). *Handbuch der physiologischen Optik*. L. Voss, Leipzig.
- Henson RN & Rugg MD (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, 41(3):263-70.
- Hesselmann V, Sorger B, Lasek K, Guntinas-Lichius O, Krug B, Sturm V, Goebel R & Lackner K (2004). Discriminating the cortical representation sites of tongue and lip movement by functional MRI. *Brain Topography*, 16:159-167.
- Hickok G, Houde J & Rong F (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, 69(3):407-22.
- Hickok G (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience* 13, 135-145
- Hickok G & Poeppel D (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Science*, 4(4):131-138.
- Hickok G & Poeppel D (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92:67-99.
- Hickok G & Poeppel D (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8:393-402.
- Hickok G, Costanzo M, Capasso R & Miceli G (2011). The role of Broca's area in speech perception: evidence from aphasia revisited. *Brain and Language*, 119(3):214-220.
- Hillis AE, Work M, Barker PB, Jacobs MA, Breese EL & Maurer K (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain*, 127(Pt 7):1479-87.
- Hodzic A, Muckli L, Singer W & Stirn A (2009). Cortical responses to self and others. *Human Brain Mapping*, 30(3):951-962
- Houde JF & Jordan MI (1998). Sensorimotor adaptation in speech production. *Science*, 279(5354):1213-6.
- Houde JF, Nagarajan SS, Sekihara K & Merzenich MM (2002). Modulation of the auditory cortex during speech: an MEG study. *Journal of Cognitive Neuroscience*, 14:1125-1138.
- Houk JC, Bastianen C, Fansler D, Fishbach A, Fraser D, Reber PJ, Roy SA & Simo LS (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(1485):1573-83.
- Humbert IA & Robbins J (2007). Normal swallowing and functional magnetic resonance imaging: a systematic review. *Dysphagia*, 22(3):266-75.
- Iwata K, Yagi J, Tsuboi Y, Koshikawa N, Sumino R & Cools AR (1996). Anatomical connections of the

- ventral, but not the dorsal part of the striatum with the parvicellular reticular formation: Implications for the anatomical substrate of oral movements. *Neuroscience Research Communications*, 18(2):71–77
- Jeannerod M (2001). Neural simulation of action: a unifying mechanism for motor cognition. *NeuroImage*, 14(1 Pt 2):S103-9.
- Jeannerod M (2003). The mechanism of self-recognition in humans. *Behavioral Brain Research*, 142(1-2):1-15. Review.
- Jeannerod, M (1997). *The cognitive neuroscience of action*. Blackwell Publishers.
- Jeannerod M (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17, 187-245.
- Jeannerod M & Jacob P (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, 9, 21-25.
- Joanisse MF, Zevin JD & McCandliss BD (2007). Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and a short-interval habituation trial paradigm. *Cerebral Cortex*, 17(9):2084-93.
- Johnson-Frey SH, Newman-Norlund R & Grafton ST (2005). A distributed left hemisphere network active during planning of everyday tool use skills. *Cerebral Cortex*, 15(6):681-95.
- Johnson-Frey SH (2004). The neural bases of complex tool use in humans. *Trends in Cognitive Sciences*, 8(2):71-8. Review.
- Jones JA & Munhall KG (2000). Perceptual calibration of F0 production: evidence from feedback perturbation. *Journal of the Acoustical Society of America*, 108(3 Pt 1):1246-51.
- Jones JA & Callan DE (2003). Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. *Neuroreport*, 14(8):1129-1133.
- Jones JA & Munhall KG (2005). Remapping auditory-motor representations in voice production. *Current Biology*, 15(19):1768-72.
- Jürgens U (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, 26(2):235-58.
- Jürgens U (2009). The neural control of vocalization in mammals, a review. *Journal of Voice*, 23(1):1-10.
- Kaas JH (2004). Somatosensory System. In G. Paxinos and J.K. Mai (Eds.), *The Human Nervous System*, 2nd ed., New York: Elsevier Academic Press, 1059-1092.
- Takei S, Hoffman DS & Strick PL (1999). Muscle and movement representations in the primary motor cortex. *Science*, 285, 2136–2139.
- Kaplan JT, Aziz-Zadeh L, Uddin L & Iacoboni M (2008). The self across the senses: The neural response to one's own face and voice. *Social, Cognitive, and Affective Neuroscience*, 3(3), pp. 218-223.
- Kappes J, Baumgaertner A, Peschke C & Ziegler W (2009). Unintended imitation in nonword repetition. *Brain and Language*, 111(3):140-51.
- Kawato M (1999). Internal models for motor control and trajectory planning. *Current opinion in neurobiology*, 9(6):718-27.
- Kawato M, Furukawa K, & Suzuki R (1987). A hierarchical neural network model for the control and learning of voluntary movements. *Biological Cybernetics*, 56, 1-17.
- Kell CA, Morillon B, Kouneiher F & Giraud AL (2011). Lateralization of speech production starts in sensory cortices--a possible sensory origin of cerebral left dominance for speech. *Cerebral Cortex*. 21(4):932-7.
- Kerzel D & Bekkering H (2000). Motor activation from visible speech: Evidence from stimulus response compatibility. *Journal of Experimental Psychology: Human Perception and Performance*.
- Keysers C, Kohler E, Umiltà MA, Fogassi L, Gallese V & Rizzolatti G (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153:628-636.

- Kilner JM, Friston KJ & Frith CD (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, 8(3):159-66.
- Kilner JM, Neal A, Weiskopf N, Friston KJ & Frith CD (2009). Evidence of mirror neurons in human inferior frontal gyrus. *Journal of Neuroscience*, 29(32):10153-9.
- Kluender KR, Diehl RL & Killeen PR (1987). Japanese Quail can learn phonetic categories. *Science*, 237:1195-1197.
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V & Rizzolatti G (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297:846-848.
- Kok P, Rahnev D, Jehee JF, Lau HC & de Lange FP (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, 22(9):2197-206.
- Kotz SA, Schwartze M & Schmidt-Kassow M (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*, 45(8), 982-990.
- Kourtzi Z & Kanwisher N (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, 293(5534):1506-9.
- Kuhl PK & Miller JD (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, 190, 69-72.
- Kuhl PK & Miller JD (1978). Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America*, 63, 905-917.
- Kuhl PK & Padden DM (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception & Psychophysics*, 32, 542-550.
- Kuhl PK (2000). A new view of language acquisition. *Proceedings of the National Academy of Science*, 97, 11850-11857.
- Ladefoged P (2006). *A Course in Phonetics*, fifth edition, 2006. Boston, MA: Thomson Wadsworth.
- Lamm C, Fischer MH & Decety J (2007). Predicting the actions of others taps into one's own somatosensory representations--a functional MRI study. *Neuropsychologia*, 45(11):2480-91.
- Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA & Fox PT (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, 10(3):120-31.
- Lane H, Denny M, Guenther FH, Hanson HM, Marrone N, Matthies ML, Perkell JS, Stockmann E, Tiede M, Vick J & Zandipour M (2007). On the structure of phoneme categories in listeners with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 50(1):2-14.
- Lane HL & Tranel B (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, 14, 677-709.
- Lauter JL, Herscovitch P, Formby C & Raichle ME (1985). Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hearing Research*, 20, 199-205.
- Leopold NA & Daniels SK (2010). Supranuclear Control of Swallowing, *Dysphagia*, 25(3):250-7.
- Lestou V, Pollick FE & Kourtzi Z (2008). Neural substrates for action understanding at different description levels in the human brain. *Journal of Cognitive Neuroscience*, 20(2):324-41.
- Leung MT & Ciocca V (2011). The effects of tongue loading and auditory feedback on vowel production. *Journal of the Acoustical Society of America*, 129(1):316-25.
- Liberman AM, Cooper FS, Harris KS & MacNeilage PF (1962). A motor theory of speech perception. *Proceedings of the Speech Communication Seminar*, Stockholm.
- Liberman AM, Cooper FS, Shankweiler DP & Studdert-Kennedy M (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.
- Liberman AM & Whalen DH (2000). On the relation of speech to language. *Trends in Cognitive Science*, 3(7):254-264.
- Liberman AM & Mattingly IG (1985). The motor theory of speech perception revised. *Cognition*, 21:1-

- Lieberman AM, Harris KS, Hoffman HS & Griffith BC (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358–368
- Lieberman AM, Isenberg D & Rakerd B (1981). Duplex perception of cues for stop consonants: Evidence for a phonetic mode. *Perception and Psychophysics*, 30, 133–143
- Lieberman AM, Mattingly IG & Turvey M (1972). Language codes and memory codes. In A.W. Melton and E. Martin (Eds.), *Coding Processes and Human Memory*. Washington, DC: Winston.
- Liepmann H (1988). Apraxie, Originally Published in *Ergebnisse der Gesamten Medizin* 1, 516–543, 1920. In J. S. Brown (Ed.), *Agnosia and apraxia: Selected papers of Liepmann, Lange and Potzl*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Liljencrants J & Lindblom B (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language*, 48, 839–862.
- Lingnau A, Gesierich B & Caramazza A (2009). Asymmetric fMRI Adaptation Reveals No Evidence for Mirror Neurons in Humans. *Proceedings of the National Academy of Sciences USA*, 106, 9925–9930.
- Liu JZ, Shan ZY, Zhang LD, Sahgal V, Brown RW & Yue GH (2003). Human brain activation during sustained and intermittent submaximal fatigue muscle contractions, an FMRI study. *Journal of Neurophysiology*, 90(1):300–12.
- Logothetis NK, Pauls J, Augath M, Trinath T & Oeltermann A (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843):150–7.
- Lombard E (1911). Le signe de l'élévation de la voix. *Annales des maladies de l'oreille et du larynx*, 37:101–119.
- Lotto A, Hickok G & Holt L (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, 13, 110–114.
- Lotto AJ, Kluender KR & Holt LL (1997). Perceptual compensation for coarticulation by Japanese quail (*Coturnix coturnix japonica*). *Journal of the Acoustical Society of America*, 102(2 Pt 1):1134–40.
- Lotze M, Seggewies G, Erb M, Grodd W & Birbaumer N (2000a). The representation of articulation in the primary sensorimotor cortex. *Neuroreport*, 11:2985–2989
- Loucks TMJ, Poletto CJ, Simonyan K, Reynolds CL & Ludlow CL (2007). Human brain activation during phonation and exhalation: common volitional Control for two upper airway functions. *NeuroImage*, 36(1):131–43.
- Lund JP & Kolta A (2006). Brainstem circuits that control mastication: do they have anything to say during speech? *Journal of Communication Disorders*, 39(5):381–90
- Macmillan NA & Creelman CD (1991). *Detection Theory: A User's Guide*. New York : Cambridge University
- MacNeillage PF (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21:499–546.
- Majdandzic J, Bekkering H, van Schie HT & Toni I (2009). Movement-specific repetition suppression in ventral and dorsal premotor cortex during action observation. *Cerebral Cortex*, 19(11):2736–45.
- Malson L (1964). *Les enfants sauvages, mythe et réalité*. Editions 10/18.
- Martin R, Barr A, MacIntosh B, Smith R, Stevens T, Taves D, Gati J, Menon R & Hachinski V (2007). Cerebral cortical processing of swallowing in older adults. *Experimental Brain Research*, 176(1):12–22.
- Martin RE, Goodyear BG, Gati JS & Menon RS (2001). Cerebral cortical representation of automatic and volitional swallowing in humans. *Journal of Neurophysiology*, 85(2):938–50.
- Martin RE, MacIntosh BJ, Smith RC, Barr AM, Stevens TK, Gati JS & Menon RS (2004). Cerebral areas processing swallowing and tongue movement are overlapping but distinct: a functional magnetic resonance imaging study. *Journal of Neurophysiology*, 92(4):2428–43.

- Massaro DM & Chen TH (2008). The motor theory of speech perception revisited. *Psychonomic Bulletin & Review*, 15(2):453-7; discussion 458-62.
- Massaro DW (1998). Categorical Perception: Important phenomenon or Lasting Myth? In R.H. Mannell & J. Robert-Ribes (Eds.), *Proceedings of the 5th International Congress of Spoken Language Processing* (pp. 2275-2279, Vol. 6). Sydney, Australia.
- Mather G & Murdoch L (1998). Evidence for global motion interactions between first-order and second-order stimuli. *Perception*, 27(7):761-7.
- Matthies ML, Guenther FH, Denny M, Perkell JS, Burton E, Vick J, Lane H, Tiede M & Zandipour M (2008). Perception and production of r allophones improve with hearing from a cochlear implant. *Journal of the Acoustical Society of America*, 124(5):3191-202.
- Mazoyer B, Zago L, Mellet E, Bricogne S, Etard O, Houdé O, Crivello F, Joliot M, Petit L & Tzourio-Mazoyer N (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, 54(3):287-98.
- McFarland DH (2001). Respiratory markers of conversational interaction. *Journal of Speech, Language, and Hearing Research*, 44:128-143
- McKay LC, Evans KC, Frackowiak RS & Corfield DR (2003). Neural correlates of voluntary breathing in humans. *Journal of Applied Physiology*, 95(3):1170-8.
- McKiernan KA, D'Angelo BR, Kaufman JN & Binder JR (2006). Interrupting the "stream of consciousness": an fMRI investigation. *NeuroImage*, 29(4):1185-91
- McKiernan KA, Kaufman JN, Kucera-Thompson J & Binder JR (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience*, 15(3):394-408.
- Meister IG, Wilson SM, Deblieck C, Wu AD & Iacoboni M (2007). The Essential Role of Premotor Cortex in Speech Perception. *Current Biology*, 17(19):1692-1696.
- Menard L, Schwartz JL & Aubin J (2008). Invariance and variability in the production of the height feature in French vowels. *Speech communication*, 50(1):14-28
- Menenti L, Gierhan SM, Segaert K & Hagoort P (2011). Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science*, 22(9):1173-82
- Merzenich MM & Brugge JF (1973). Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain Research*, 50:275-296.
- Meyer M (2008). Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. *Zeitschrift für Neuropsychologie*, 19, 101-114.
- Miceli G, Gainotti G, Caltagirone C & Masullo C (1980). Some aspects of phonological impairment in aphasia. *Brain and Language*, 11:159-169.
- Mielke J, Baker A, & Archangeli D (2006). Covert /ɹ/ allophony in English: variation in a socially uninhibited sound pattern. Paper presented at *Laboratory Phonology 10*, Paris.
- Mishkin M & Ungerleider LG (1982). Object Vision and Spatial Vision: Two Cortical Pathways. In DJ Ingle, MA Goodale, and RJW Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, MA: MIT Press 296-302
- Morel A, Garraghty PE & Kaas JH (1993). Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *The Journal of Comparative Neurology*, 335:437-459.
- Möttönen R & Watkins KE (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *The Journal of Neuroscience*, 29(31):9819-9825.
- Möttönen R, Järveläinen J, Sams M & Hari R (2004). Viewing speech modulates activity in the left SI mouth cortex. *NeuroImage*, 24:731-737.
- Munhall KG, MacDonald EN, Byrne SK & Johnsrude I (2009). Talkers alter vowel production in

- response to real-time formant perturbation even when instructed not to compensate. *Journal of the Acoustical Society of America*, 125(1):384-90.
- Murphy K, Corfield DR, Guz A, Fink GR, Wise RJ, Harrison J & Adams L (1997). Cerebral areas associated with motor control of speech in humans. *Journal of Applied Physiology*, 83(5):1438-47.
- Mussa-Ivaldi FA (1999). Modular features of motor control and learning. *Current opinion in neurobiology*, 9(6):713-7.
- Nakamura K, Kawashima R, Sugiura M, Kato T, Nakamura A, Hatano K, Nagumo S, Kubota K, Fukuda H, Ito K & Kojima S (2001). Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia*, 39(10):1047-54.
- Nakamura Y & Katakura N (1995). Generation of masticatory rhythm in the brainstem. *Neuroscience Research*, 23(1):1-19
- Nakano K, Tokushige A, Kohno M, Hasegawa Y, Kayahara T & Sasaki K (1992). An autoradiographic study of cortical projections from motor thalamic nuclei in the macaque monkey. *Neuroscience Research*, 13:119-137.
- Nasir SM & Ostry DJ (2006). Somatosensory precision in speech production. *Current Biology*, 16:1918–1923.
- Natale M (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, 32:790-804.
- Nelissen K, Luppino G, Vanduffel W, Rizzolatti G & Orban GA (2005). Observing others: multiple action representation in the frontal lobe. *Science*, 14;310(5746):332-6.
- Nichols T, Brett M, Andersson J, Wager T & Poline JB (2005). Valid conjunction inference with the minimum statistic. *NeuroImage*, 25(3):653-60.
- Nichols T & Hayasaka S (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical Methods in Medical Research*, 12(5):419-46.
- Nikouline V, Ruohonen J & Ilmoniemi RJ (1999). The role of the coil click in TMS assessed with simultaneous EEG. *Clinical Neurophysiology*, 110:1325-1328.
- Nishitani N & Hari R (2002). Viewing lip forms: Cortical dynamics. *Neuron*, 36:1211-1220.
- Nixon P, Lazarova J, Hodinott-Hill I, Gough P & Passingham R (2004). The inferior frontal gyrus and phonological processing: an investigation using rTMS. *Journal of Cognitive Neuroscience*, 16(2):289-300.
- Norman KA, Polyn SM, Detre GJ & Haxby JV (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9):424-30.
- Nota Y & Honda K (2004). Brain regions involved in motor control of speech. *Acoustical Science and Technology*, 25(4):286-289.
- Numminen J & Curio G (1999). Differential effects of overt, covert and replayed speech on vowel-evoked responses of the human auditory cortex. *Neuroscience Letters*, 272:29–32.
- Obeso JA, Rodríguez-Oroz MC, Rodríguez M, Lanciego JL, Artieda J, Gonzalo N & Olanow CW (2000). Pathophysiology of the basal ganglia in Parkinson's disease. *Trends in Neurosciences*, 23:S8-S19.
- Obleser J & Eisner F (2009). Pre-lexical abstraction of speech in the auditory cortex. *Trends in Cognitive Sciences*, 13(1):14-9.
- Obleser J, Elbert T, Lahiri A & Eulitz C (2003). Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Brain Research. Cognitive Brain Research*, 15(3):207-13.
- Obleser J, Wise RJ, Alex Dresner M & Scott SK (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience*, 27(9):2283-9.
- Obleser J, Boecker H, Drzezga A, Haslinger B, Hennenlotter A, Roetinger M, Eulitz C & Rauschecker JP (2006). Vowel sound extraction in anterior superior temporal cortex. *Human Brain Mapping*,

27(7):562–571.

- Oblaser J, Leaver AM, Vanmeter J & Rauschecker JP (2010). Segregation of vowels and consonants in human auditory cortex: Evidence for distributed hierarchical organization. *Front. Psychology*, 1:232.
- Ojanen V, Möttönen R, Pekkola J, Jääskeläinen IP, Joensuu R, Autti T & Sams M (2005). Processing of audiovisual speech in Broca's area. *NeuroImage*, 25:333–338.
- Okada K & Hickok G (2006). Left posterior auditory-related cortices participate both in speech perception and speech production: Neural overlap revealed by fMRI. *Brain and Language*, 98(1):112–7.
- Oldfield RC (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9:97–114.
- Onozuka M, Hirano Y, Tachibana A, Kim W, Ono Y, Sasaguri K, Kubo K, Niwa M, Kanematsu K & Watanabe K (2007). *Interactions between chewing and brain activity in humans*. Novel Trends in Brain Science (Eds M Onozuka, CT Yen), pp. 99–113, Springer.
- Özdemir E, Norton A & Schlaug G (2006). Shared and distinct neural correlates of singing and speaking. *NeuroImage*, 33:628–635.
- Pardo J (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119(4):2382–93.
- Paulesu E, Frith C & Frackowiak R (1993). Functional anatomy of the verbal component of working memory. *Nature*, 362:342–344.
- Paulesu E, Perani D, Blasi V, Silani G, Borghese AA, De Giovanni U, Sensolo S & Fazio F (2003). A functional-anatomical model for lipreading. *Journal of Neurophysiology*, 90:2005–2013.
- Paus T, Perry D, Zatorre R, Worsley KJ & Evans AC (1996). Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. *European Journal of Neuroscience*, 8:2236–2246.
- Peereman R, Dufour S & Burt JS (2009). Orthographic influences in spoken word recognition: the consistency effect in semantic and gender categorization tasks. *Psychonomic Bulletin & Review*. 16(2):363–8.
- Peeva MG, Guenther FH, Tourville JA, Nieto-Castanon A, Anton JL, Nazarian B & Alario FX (2009). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *NeuroImage*, 50(2):626–38.
- Pekkola J, Ojanen V, Autti T, Jääskeläinen IP, Möttönen R, Tarkiainen A & Sams M (2005). Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *Neuroreport*, 6(2):125–8.
- Pekkola J, Laasonen M, Ojanen V, Autti T, Jaaskelainen LP, Kujala T & Sams M (2006). Perception of matching and conflicting audiovisual speech in dyslexic and fluent readers: an fMRI study at 3T. *NeuroImage*, 29(3):797–807.
- Pelgrims B, Andres M & Olivier E (2009). Double dissociation between motor and visual imagery in the posterior parietal cortex. *Cerebral Cortex*, 19(10):2298–307.
- Penfield W & Boldrey E (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, 60:389–443.
- Penfield W & Rasmussen T (1950). *The Cerebral Cortex of Man*. New York: Macmillan.
- Perkell JS, & Klatt DH (1986). *Invariance and variability in speech processes*. Hillsdale N.J, New Jersey: L. Erlbaum.
- Perkell JS (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics*, 25(5):382–407.
- Perkell JS, Matthies ML, Svirsky MA & Jordan MI (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot motor equivalence study. *Journal of the Acoustical Society of America*, 93:2948–2961.

- Perrier P (2005). Control and representations in speech production. *ZAS Papers in Linguistics*, 40:109-132.
- Perrier P (2012). Gesture planning integrating knowledge of the motor plant's dynamics: A literature review from motor control and speech motor control. *Speech Planning and Dynamics, Collection: Speech Production and Perception*, 191-238.
- Perrone-Bertolotti M (2011). *Spécialisation hémisphérique et cartographie du langage chez le sujet sain et le patient épileptique : effet des variables inter- et intra-individuelles. Évaluation comportementale, neuropsychologique et neurophysiologique (IRM fonctionnelle)*. Thèse de Doctorat, Grenoble Université.
- Peschke C, Ziegler W, Kappes J & Baumgaertner A (2009). Auditory-motor integration during fast repetition: the neuronal correlates of shadowing. *NeuroImage*, 47(1):392-402.
- Petrides M, Cadoret G & Mackey S (2005). Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature*. 435:1235–1238.
- Pickard N & Strick P (2001). Imaging the premotor areas. *Current opinion in neurobiology*, 11(6):663-672.
- Pinto S, Ghio A, Teston B & Viallet F (2007). Dysarthria across Parkinson's disease progression. Natural history of its components: dysphonia, dysprosody and dysarthria. *Revue Neurologique*, 166(10):800-10.
- Pisoni DB (1977). Identification and discrimination of the relative onset of two component tones: Implications for voicing perception in stop consonants. *Journal of the Acoustical Society of America*, 61:1352-1361.
- Poeppel D (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication*, 41:245-255.
- Poeppel D, Idsardi WJ & van Wassenhove V (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 363(1493):1071-86.
- Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, & Gabrieli JD (1999). Functional specialization for semantic and phonological processing in the left inferiorprefrontal cortex. *NeuroImage*, 10:15-35.
- Pons TP, Garraghty PE, Cusick CG & Kaas JH (1985). The somatotopic organization of area 2 in macaque monkeys. *The Journal of Comparative Neurology*, 241:445--466.
- Porter RJ & Castellanos FX (1980). Speech-production measures of speech perception: rapid shadowing of VCV syllables. *Journal of Acoustical Society of America*, 67:1349–1356.
- Porter RJ Jr. & Lubker JF (1980). Rapid reproduction of vowel–vowel sequences: evidence for a fast and direct acousticmotoric linkage in speech. *Journal of Speech and Hearing Research*, 23:593–602.
- Price C (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*. 1191:62-88.
- Price C, Thierry G & Griffiths T (2005). Speech-specific auditory processing: where is it? *Trends in Cognitive Sciences*. 9(6):271-6.
- Price CJ, Crinion JT & MacSweeney M (2011). A generative model of speech production in Broca's and Wernicke's areas. *Frontiers in Psychology*, 2:237.
- Pulvermüller F, Huss M, Kherif F, Moscoso del Prado Martin F, Hauk O & Shtyrov Y (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the USA*, 103(20):7865-70.
- Purcell DW & Munhall KG (2006). Adaptive control of vowel formant frequency: evidence from real-time formant manipulation. *Journal of the Acoustical Society of America*. 120(2):966-77.
- Purves D, Augustine GJ, Fitzpatrick D, Katz LC, LaMantia A-S & McNamara (2005). *Neurosciences*

- (traduction J-M Coquery), éditions DeBoeck.
- Rahnev D, Lau H & de Lange FP (2011). Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. *Journal of Neuroscience*, 2011 Jul 20;31(29):10741-8.
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA & Shulman GL (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the USA*, 98:676–682.
- Raij T (1999). Patterns of brain activity during visual imagery of letters. *Journal of Cognitive Neuroscience*, 11(3):282-99.
- Raizada RD & Poldrack RA (2007). Selective amplification of stimulus differences during categorical processing of speech. *Neuron*, 56(4):726-40.
- Rand TC (1974). Dichotic release of masking speech. *Journal of the Acoustical Society of America*, 55(3):678-80.
- Rapin L (2011). *Hallucinations auditives verbales et trouble du langage intérieur dans la schizophrénie : traces physiologiques et bases cérébrales*. Thèse de Doctorat, Grenoble Université.
- Rauschecker JP & Scott SK (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6):718–724.
- Rauschecker JP (2012). Ventral and dorsal streams in the evolution of speech and language. *Frontiers in Evolutionary Neuroscience*, 4:7.
- Rauschecker JP, Tian B & Hauser M (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, 268(5207):111-4.
- Rauschecker JP (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, 271:16-25.
- Reina GA, Moran DW & Schwartz AB (2001). On the relationship between joint angular velocity and motor cortical discharge during reaching. *Journal of Neurophysiology*, 85(6):2576-89.
- Riecker A, Ackermann H, Wildgruber D, Dogil G & Grodd W (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport*, 11(9):1997-2000.
- Riecker A, Ackermann H, Wildgruber D, Meyer J, Dogil G, Haider H & Grodd W (2000). Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: a functional magnetic resonance imaging (fMRI) study. *Brain and Language*, 75(2):259-76.
- Riecker A, Mathiak K, Wildgruber D, Erb M, Hertrich I, Grodd W & Ackermann H (2005). fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology*, 64(4):700-6.
- Rizzolatti G, Camarda R, Fogassi L, Gentilucci M, Luppino G & Matelli M (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research*, 71:491–507.
- Rizzolatti G & Craighero L (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27:169-92.
- Rizzolatti G & Fabbri-Destro M (2008). The mirror system and its role in social cognition. *Current Opinion in Neurobiology*, 18(2):179-84
- Rizzolatti G, Fadiga L, Gallese V & Fogassi L (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3:131-141.
- Rizzolatti G, Fogassi L & Gallese V (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature reviews. Neuroscience*, 2:661-670.
- Rizzolatti G & Sinigaglia C (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature reviews. Neuroscience*, 11(4):264-74.
- Rizzolatti G & Arbib MA (1998). Language within our grasp. *Trends in Neurosciences*, 21:188-194.
- Rizzolatti G & Craighero L (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27:169-192.
- Romero, L., Walsh, V. & Papagno, C (2006). The neural correlates of phonological short-term

- memory: a repetitive transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 18(7):1147-1155.
- Rosa C, Lassonde M, Pinard C, Keenan JP & Belin P (2008). Investigations of hemispheric specialization of self-voice recognition. *Brain Cognition*, 68(2):204-14.
- Rossini PM, Barker AT, Berardelli A, Caramia MD, Caruso G, Cracco RQ, Dimitrijevic MR, Hallett M, Katayama Y, Lücking CH, et al (1994). Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalography and Clinical Neurophysiology*, 91(2):79–92.
- Roy AC, Craighero L, Fabbri-Destro M & Fadiga L (2008). Phonological and lexical motor facilitation during speech listening: A transcranial magnetic stimulation study. *Journal of physiology*, Paris, 102(1-3):101-105.
- Rozzi S, Ferrari PF, Bonini L, Rizzolatti G & Fogassi L (2008). Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *European Journal of Neuroscience*, 28(8):1569-88.
- Sakata H, Taira M, Murata A & Mine S (1995). Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cerebral Cortex*, 5(5):429-38.
- Saltzman, E (1986). Task dynamic coordination of the speech articulators: A preliminary model. In H. Heuer & C. Fromm (Eds.), *Experimental Brain Research Series 15* (pp. 129-144). New York: Springer-Verlag.
- Samuel AG (1986). Red herring detectors and speech perception: In defense of selective adaptation. *Cognitive Psychology*, 18:452-499.
- Sancier ML & Fowler CA (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25:421-436.
- Sato M, Buccino G, Gentilucci M & Cattaneo L (2010). On the tip of the tongue: modulation of the primary motor cortex during audiovisual speech perception. *Speech Communication*, 52(6):533-541.
- Sato M (2004). *Représentations verbales multistables en mémoire de travail : Vers une perception active des unités de parole*. Thèse de Doctorat, Grenoble Université.
- Sato M, Grabski K, Glenberg A, Brisebois A, Basirat A, Ménard L & Cattaneo L (2011). Articulatory bias in speech perception: evidence from use-induced motor plasticity. *Cortex*, 47(8):1001-3.
- Sato M, Tremblay P & Gracco V (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language*, 111(1):1-7.
- Savariaux C, Perrier P & Orliaguet JP (1995). Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: a study of the control space in speech production. *Journal of the Acoustical Society of America*, 98:2428-2842.
- Sawczuk A & Mosier KM (2001). Neural control of tongue movement with respect to respiration and swallowing. *Critical Reviews in Oral Biology and Medicine*, 12(1):18-37.
- Schäring M, Idsardi WJ & Poe S (2011). A comprehensive three-dimensional cortical map of vowel space. *Journal of Cognitive Neuroscience*, 23(12):3972-82.
- Schulz GM, Varga M, Jeffries K, Ludlow CL & Braun AR (2005). Functional neuroanatomy of human vocalization: an H215O PET study. *Cerebral Cortex*, 15(12):1835-47.
- Schwartz JL, Sato M & Fadiga L (2008). The common language of speech perception and action: a neurocognitive perspective. *Revue Française de Linguistique Appliquée*, 13(2):9-22.
- Schwartz JL, Abry C, Boë LJ & Cathiard MA (2002). Phonology in a theory of perception-for-action-control. In Durand, J., Lacks, B (Eds.), *Phonology: From Phonetics to Cognition*. Oxford University Press, Oxford, pp. 240-280.
- Schwartz JL, Boë LJ, Vallée N & Abry C (1997a). Major trends in vowel system inventories. *Journal of*

- Phonetics*, 25:233-254.
- Schwartz JL, Boë LJ, Vallée N & Abry C (1997b). The dispersion-focalization theory of vowel systems. *Journal of Phonetics*, 25:255-286.
- Schwartz JL, Ménard L, Basirat A & Sato M (2012). The Perception for Action Control Theory (PACT): a perceptuo-motor theory of speech perception. *Journal of Neurolinguistics*, 25(5):336-354.
- Scott SH & Kalaska JF (1997). Reaching movements with similar hand paths but different arm orientation. I. Activity of individual cells in motor cortex. *Journal of Neurophysiology*, 77: 826–852.
- Scott SH, Sergio LE & Kalaska JF (1997). Reaching movements with similar hand paths but different arm orientations. II. Activity of individual cells in dorsal premotor cortex and parietal area 5. *Journal of Neurophysiology*, 78:2413–2426.
- Scott SK & Wise RJ (2004). The functional neuroanatomy of prelexical processing in speech perception. *Cognition*, 92(1-2):13-45.
- Scott SK & Johnsrude IS (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, 26(2):100–107.
- Scott SK, McGettigan C & Eisner F (2009). A little more conversation, a little less action--candidate roles for the motor cortex in speech perception. *Nature reviews. Neuroscience*, 10(4):295-302.
- Seger CA & Cincotta CM (2005). The roles of the caudate nucleus in human classification learning. *Journal of Neuroscience*, 25(11):2941-51.
- Seger CA (2006). The basal ganglia in human learning. *Neuroscientist*. 12(4):285-90.
- Serniclaes W, Sprenger-Chraolles L, Carré R & Démonet JF (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech, Language and Hearing Research*, 44(2):384-399.
- Sessle BJ, Adachi K, Avivi-Arber L, Lee J, Nishiura H, Yao D & Yoshino K (2007). Neuroplasticity of face primary motor cortex control of orofacial movements. *Archives of Oral Biology*, 52(4):334-7.
- Shestakova A, Brattico E, Soloviev A, Klucharev V & Huotilainen M (2004). Orderly cortical representation of vowel categories presented by multiple exemplars. *Brain Research. Cognitive Brain Research*, 21(3):342-50.
- Shiller D, Sato M, Gracco V & Baum S (2009). Perceptual recalibration of speech sounds following speech motor learning. *Journal of the Acoustical Society of America*, 125(2):1103-113.
- Shockley K, Santana MV & Fowler CA (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29:326-332.
- Shum M, Shiller DM, Baum SR & Gracco VL (2011). Sensorimotor integration for speech motor learning involves the inferior parietal cortex. *European Journal of Neuroscience*, 34(11):1817-22.
- Silvanto J, Muggleton N & Walsh V (2008). State-dependency in brain stimulation studies of perception and cognition. *Trends in Cognitive Sciences*, 12:447-454.
- Simonyan K, Ostuni J, Ludlow CL & Horwitz B (2009). Functional but not structural networks of the human laryngeal motor cortex show left hemispheric lateralization during syllable but not breathing production. *Journal of Neuroscience*, 29(47):14912-23.
- Skipper JI, Nusbaum HC & Small SL (2005). Listening to talking faces: Motor cortical activation during speech perception. *NeuroImage*, 25:76-89.
- Skipper JI, Van Wassenhove V, Nusbaum HC & Small SL (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, 17(10):2387-2399.
- Smith A (1998). The control of orofacial movements in speech. *Critical Reviews in Oral Biology and Medicine*, 3(3):233-67
- Smotherman MS (2007). Sensory feedback control of mammalian vocalizations. *Behavioural Brain Research*, 182(2):315-26

- Sörös P, Sokoloff LG, Bose A, McIntosh AR, Graham SJ & Stuss DT (2006). Clustered functional MRI of overt speech production. *NeuroImage*, 32(1):376-387.
- Sowman PF, Flavel SC, McShane CL, Sakuma S, Miles TS & Nordstrom MA (2009). Asymmetric activation of motor cortex controlling human anterior digastric muscles during speech and target-directed jaw movements. *Journal of Neurophysiology*, 102(1):159-66.
- Sperry RW (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, 43:482-489.
- Stevens KN & Blumstein SE (1978). Invariant cues for place of articulation in stop consonants., *Journal of the Acoustical Society of America*, 64:1358-68.
- Stevens KN & Klatt DH (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America*, 55:653-59
- Stevens KN (1972). The quantal nature of speech: evidence from articulatory-acoustic data. In E. E. Davis, Jr., & P. B. Denes (Eds.), *Human communication: A unified view* (pp. 51-66). New-York: Mc Graw-Hill.
- Stevens KN (1989). On the quantal nature of speech. *Journal of Phonetics*, 17:3-45.
- Stevens KN & Halle M (1967). Remarks on analysis by synthesis and distinctive features. In W. Walthe-Dunn (Ed.), *Models for the perception of speech and visual form*. Cambridge, MA: MIT Press, pp. 88-102.
- Sugiura M, Watanabe J, Maeda Y, Matsue Y, Fukuda H & Kawashima R (2005). Cortical mechanisms of visual self-recognition. *NeuroImage*, 24(1):143-9.
- Summerfield C, Egner T, Greene M, Koechlin E, Mangels J & Hirsch J (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science*, 314(5803):1311-4.
- Summerfield C, Trittschuh EH, Monti JM, Mesulam MM & Egner T (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, 11(9):1004-6.
- Sundara M, Namasivayam AK & Chen R (2001). Observation-execution matching system for speech: A magnetic stimulation study. *Neuroreport*, 12(7):1341-1344.
- Svensson P & Romaniello A (2003). Plasticity in corticomotor control of the human tongue musculature induced by tongue-task training. *Experimental Brain Research*, 152:42-51.
- Takai O, Brown S & Liotti M (2010). Representation of the speech effectors in the human motor cortex: somatotopy or overlap? *Brain and Language*, 113:39-44.
- Takaso H, Eisner F, Wise RJ & Scott SK (2010). The effect of delayed auditory feedback on activity in the temporal lobe while speaking: a positron emission tomography study. *Journal of Speech, Language, and Hearing Research*, 53(2):226-36.
- Talairach J & Tournoux P (1988). Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishers.
- Teichmann M, Darcy I, Bachoud-Lévi AC & Dupoux E (2009). The role of the striatum in phonological processing: evidence from early stages of Huntington's disease. *Cortex*, 45(7):839-49.
- Terumitsu M, Fujii Y, Suzuki K, Kwee IL & Nakada T (2006). Human primary motor cortex shows hemispheric specialization for speech. *Neuroreport*, 17(11):1091-5.
- Thach WT (1998). A role for the cerebellum in learning movement coordination. *Neurobiology of Learning and Memory*, 70(1-2):177-88.
- Tian X & Poeppel D (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, 1:166.
- Todorovic A, van Ede F, Maris E & de Lange FP (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. *Journal of Neuroscience*, 31(25):9118-23.
- Tourville JA, Reilly KJ & Guenther FH (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, 39(3):1429-43.

- Toyomura A, Koyama S, Miyamaoto T, Terao A, Omori T, Murohashi H & Kuriki S (2007). Neural correlates of auditory feedback control in human. *Neuroscience*, 146(2):499-503.
- Tremblay P & Small SL (2011). On the context-dependent nature of the contribution of the ventral premotor cortex to speech perception. *NeuroImage*, 57(4):1561-71.
- Tremblay S, Shiller DM & Ostry DJ (2003). Somatosensory basis of speech production. *Nature*, 423(6942):866-9.
- Turkeltaub PE & Coslett HB (2010). Localization of sublexical speech perception components. *Brain and Language*, 114(1):1-15.
- Uddin LQ, Kaplan JT, Molnar-Szakacs I, Zaidel E & Iacoboni M (2005). Self-face recognition activates a frontoparietal "mirror" network in the right hemisphere: an event-related fMRI study. *NeuroImage*, 25(3):926-35.
- Uddin LQ, Molnar-Szakacs I, Zaidel E & Iacoboni M (2006). rTMS to the right inferior parietal lobule disrupts self-other discrimination. *Social Cognitive and Affective Neuroscience*, 1(1):65-71.
- Umiltà MA, Brochier T, Spinks RL & Lemon RN (2007). Simultaneous recording of macaque premotor and primary motor cortex neuronal populations reveals different functional contributions to visuomotor grasp. *Journal of Neurophysiology*, 98(1):488-501.
- Umiltà MA, Kohler E, Gallese V, Fogassi L, Fadiga L, Keysers C & Rizzolatti G (2001). I Know What You Are Doing: A Neurophysiological Study. *Neuron*, 31(1):155-65.
- Uppenkamp S, Johnsrude IS, Norris, D., Marslen-Wilson, W. & Patterson, R.D (2006). Locating the initial stages of speech-sound processing in human temporal cortex. *NeuroImage*, 31:1284-1296.
- Vaden KI Jr, Muftuler LT & Hickok G (2010). Phonological repetition-suppression in bilateral superior temporal sulci. *NeuroImage*, 49(1):1018-23.
- Valentini E (2011). The role of perceptual expectation on repetition suppression, a quest to dissect the differential contribution of probability of occurrence and event predictability. *Frontiers in Human Neuroscience*, 5:143.
- Vallée N (1994). *Systèmes vocaliques : de la typologie aux prédictions*. Thèse de Doctorat, Université Stendhal, Grenoble.
- vanDuinen H, Renken R, Maurits N & Zijdewind I (2007). Effects of motor fatigue on human brain activity, an fMRI study. *NeuroImage*, 35(4):1438-49
- Vanzetta I & Grinvald A (2008). Coupling between neuronal activity and microcirculation: implications for functional brain imaging. *HFSP Journal*, 2(2):79-98
- Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O, Mazoyer B & Tzourio-Mazoyer N (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*, 30:1414-1432.
- Villacorta VM, Perkell JS & Guenther FH (2007). Sensorimotor adaptation to feedback perturbations on vowel acoustics and its relation to perception. *Journal of Acoustical Society of America*, 122:2306-2319.
- Viviani P & Stucchi N (1992). Biological movements look uniform: evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance*, 18:603-623.
- Vogt C & Vogt O (1919). Ergebnisse unserer Hirnforschung. 1.-4. Mitteilung. *Journal für Psychologie und Neurologie*, 25:1:361-376.
- Vogt C & Vogt O (1926). Die vergleichend-architektonische und die vergleichend-reizphysiologische Felderung der Grosshirnrinde unter besonderer Berücksichtigung der menschlichen. *Naturwissenschaften*, 14:1190-1194.
- von Holst, E (1954). Relations between the central Nervous System and the peripheral organs. *The British Journal of Animal Behaviour*, 2 (3):89-86.
- Watanabe J, Sugiura M, Miura N, Watanabe Y, Maeda Y, Matsue Y & Kawashima R (2004). The

- human parietal cortex is involved in spatial processing of tongue movement-an fMRI study. *NeuroImage*, 21(4):1289-99.
- Watkins KE & Paus T (2004). Modulation of motor excitability during speech perception: the role of Broca's area. *Journal of Cognitive Neuroscience*, 16(6):978-987.
- Watkins KE, Strafella AP & Paus T (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41(3):989-994.
- Webster MA, Kaping D, Mizokami Y & Duhamel P (2004). Adaptation to natural facial categories. *Nature*, 428(6982):557-61.
- Wessinger CM, Buonocore MH, Kussmaul CL & Mangun GR (1997). Tonotopy in human auditory cortex examined with functional magnetic resonance imaging. *Human Brain Mapping*, 5(1):18-25.
- Wilson SM & Iacoboni M (2006). Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. *NeuroImage*, 33(1):316-25.
- Wilson SM, Saygin AP, Sereno MI & Iacoboni M (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7:701-702.
- Wise RJ, Greene J, Büchel C & Scott SK (1999). Brain regions involved in articulation. *Lancet*, 353(9158):1057-61.
- Wolpert DM & Flanagan JR (2001). Motor prediction. *Current Biology*, 18:R729-R732.
- Wolpert DM & Ghahramani Z (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, 3 Suppl:1212-7.
- Wolpert DM, Ghahramani Z & Jordan MI (1995). An internal model for sensorimotor integration. *Science*, 269:1880-1882.
- Wolpert DM & Kawato M (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 1(7-8):1317-29.
- Wolpert DM, Miall RC & Kawato M (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9):338-47.
- Wolpert DM (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, 1(6): 209-216.
- Zaehle T, Schmidt CF, Meyer M, Baumann S, Baltes C, Boesiger P & Jancke L (2007). Comparison of "silent" clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. *NeuroImage*, 37(4):1195-204.
- Zatorre RJ & Belin P (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11,946–953.
- Zatorre RJ, Belin P & Penhune VB (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, 6(1):37-46.
- Zatorre RJ, Evans AC, Meyer E & Gjedde A (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256(5058):846-9.
- Zatorre RJ & Gandour JT (2008). Neural specializations for speech and pitch, moving beyond the dichotomies *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 363(1493):1087-104.
- Zatorre RJ (2007). There's more to auditory cortex than meets the ear. *Hearing Research*, 229(1-2):24-30.
- Zekveld AA, Heslenfeld DJ, Festen JM & Schoonhoven R (2006). Top-down and bottom-up processes in speech comprehension. *NeuroImage*, 32:1826-1836.
- Zheng ZZ, Munhall KG & Johnsrude IS (2010). Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *Journal of Cognitive Neuroscience*, 22(8):1770-81.

Liste des Publications

ARTICLES PUBLIES OU SOUMIS DANS DES REVUES AVEC COMITE DE LECTURE

- Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M (2012). Functional MRI assessment of orofacial articulators: neural correlates of lip, jaw, larynx and tongue movements. *Human Brain Mapping*, 33(10):2306-21.
- Sato M, Grabski K, Glenberg A, Brisebois A, Basirat A, Ménard L & Cattaneo L (2011). Articulatory bias in speech perception: evidence from use-induced motor plasticity. *Cortex*, 47(8):1001-3.
- Grabski K, Lamalle L & Sato M (sous presse). Somatosensory-motor adaptation of orofacial actions in left posterior parietal and ventral premotor cortices. *PlosOne*.
- Grabski K, Tremblay P, Gracco V, Girin L & Sato M (soumis). A mediating role of the auditory dorsal pathway in selective adaptation to speech: a state-dependent transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*.
- Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Baciú M, Le Bas JF & Sato M (soumis). Shared and distinct neural substrates in vowel perception and production. *Journal of Speech, Language and Hearing Research*.
- Grabski K, Lamalle L & Sato M (en préparation). Mirror mechanisms and sensorimotor adaptive changes during vowel perception and production.
- Sato M, Grabski K, Garnier M, Granjon L, Schwartz JL & Nguyen N (en préparation). Converging to a common speech code: imitative and perceptuo-motor recalibration processes in speech production.

COMMUNICATION ORALE INVITEE

- Grabski K (2011). Sensorimotor maps of speech: functional coupling and sensorimotor predictive coding of vowel perception & production. *Workshop 'Imitation and adaptation between interlocutors'*, Laboratoire Parole & Langage, Décembre 09, Aix-en Provence, France.

ACTES ET RESUMES

Réseaux cérébraux des systèmes de perception et production des voyelles du Français et des articulateurs orofaciaux

- Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M (2010). Motor somatotopy of supralaryngeal and laryngeal articulators. *16th Annual Meeting of the Organization for Human Brain Mapping*, 6-10 Juin 2010, Barcelone, Espagne.
- Grabski K, Lamalle L, Schwartz JL, Vilain C, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M (2010). Neural correlates of vowel perception and production. *16th Annual Meeting of the Organization for Human Brain Mapping*, 6-10 Juin 2010, Barcelone, Espagne.
- Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M (2010). Somatotopie motrice des articulateurs supralaryngés de la parole. *XXVIIIèmes Journées d'Étude sur la Parole*, 25-28 Mai 2010, Mons, Belgique.
- Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M (2010). Corrélats neuroanatomiques des systèmes de perception et de production des voyelles du Français. *XXVIIIèmes Journées d'Étude sur la Parole*, 25-28 Mai 2010, Mons, Belgique.

- Sato M, Grabski K, Lamalle L, Schwartz JL, Vilain C, Vallée N, Troprès I, Baciú M & Le Bas JF (2010). Shared and distinct neural correlates of vowel perception and production. *Seventeenth Annual Cognitive Neuroscience Society Meeting*, 17-20 Avril 2010, Montréal, Canada.
- Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M (2010). Somatotopic representation of laryngeal and supralaryngeal movements in the sensorimotor cortex. *Seventeenth Annual Cognitive Neuroscience Society Meeting*, 17-20 Avril 2010, Montréal, Canada.
- Grabski K, Schwartz JL, Lamalle L, Vilain C, Vallée N, Troprès I, Baciú M, Le Bas JF & Sato M (2009). Représentations cérébrales des articulateurs de la parole. *VIIIème Rencontres des Jeunes Chercheurs en Parole*, 16-18 Novembre 2009, Avignon, France.

Contrôle sensorimoteur adaptatif et codage du but des actions orofaciales et vocaliques

- Grabski K & Sato M (2012). Speech as a sensorimotor process: sensorimotor predictive coding and functional coupling of vowel perception and production. *International Symposium on Imitation and Convergence in speech*, Septembre 03-05, Aix-en-Provence, France.
- Grabski K, Lamalle L & Sato M (2012). Contrôle prédictif et codage du but des actions orofaciales. *Proceedings of the XXIXèmes Journées d'Étude sur la Parole*, Juin 04-08, Grenoble, France.
- Grabski K & Sato M (2012). Les cartes sensorimotrices de la parole : codage prédictif sensorimoteur et couplage fonctionnel des systèmes de perception et de production des voyelles. *Journée thématique du Pôle Grenoble Cognition*, Avril 02, Grenoble, France.
- Grabski K, Lamalle L & Sato M (2011). Mirror mechanisms and predictive coding during vowel perception and production. *17th Annual Meeting of the Organization for Human Brain Mapping*, Juin 26-30, Québec, Canada.
- Grabski K, Lamalle L & Sato M (2011). Functional coupling and sensorimotor predictive coding of vowel perception and production. *Ninth International Seminar on Speech Production*, Juin 20-23, Montreal, Canada.
- Grabski K, Lamalle L & Sato M (2011). Motor control and action goal coding of orofacial movements. *17th Annual Meeting of the Organization for Human Brain Mapping*, Juin 26-30, Québec, Canada.
- Grabski K, Lamalle L & Sato M (2011). Core neural network, somatotopy and predictive sensorimotor coding of orofacial movements. *Ninth International Seminar on Speech Production*, Juin 20-23, Montreal, Canada.
- Grabski K, Lamalle L & Sato M (2011). Corrélats neurocognitifs des systèmes de perception et de production de la parole : Couplage fonctionnel et codage sensorimoteur prédictif. *Xème Rencontres des Jeunes Chercheurs en Parole*, Mai 25-27, Grenoble, France.
- Grabski K, Lamalle L & Sato M (2010). Mirror mechanisms and predictive sensorimotor coding during vowel perception and production: a sparse-sampling adaptation fMRI study. *Second Neurobiology of Language Conference*, Novembre 11-12, San Diego, USA.

Mécanismes de convergence phonétique et d'imitation de la parole

- Sato M, Grabski K, Garnier L, Granjon L, Schwartz JL & Nguyen N (2012). Plasticity of auditory goals in speech production: behavioral evidence from phonetic convergence and voluntary imitation of speech. *International Symposium on Imitation and Convergence in speech*, September 03-05, Aix-en-Provence, France.
- Sato M, Grabski K, Granjon L, Schwartz JL & Nguyen N (2011). Plasticity of auditory goals in speech production: behavioral evidence from phonetic convergence and speech imitation. *Ninth International Seminar on Speech Production*, Juin 20-23, Montreal, Canada.
- Sato M, Grabski K, Granjon L, Schwartz JL & Nguyen N (2010). Converging to a common speech code: automatic imitative and perceptuo-motor recalibration processes in speech

communication. *Second Neurobiology of Language Conference*, Novembre 11-12, San Diego, USA.

Rôle causal du système moteur et de la voie dorsale lors de la perception de la parole

- Grabski K, Tremblay P, Gracco V & Sato M (2011). A mediating role of the auditory-motor dorsal pathway in selective adaptation to speech. *Third Neurobiology of Language Conference*, Novembre 10-11, Annapolis, USA.
- Grabski K, Tremblay P, Gracco V & Sato M (2011). A mediating role of temporal, parietal, and premotor cortices in selective adaptation to speech: a state-dependent transcranial magnetic stimulation study. *17th Annual Meeting of the Organization for Human Brain Mapping*, Juin 26-30, Québec, Canada.
- Sato M, Grabski K, Brisebois A, Glenberg A, Basirat A & Cattaneo L (2009). Use-induced motor plasticity affects speech perception. *First Neurobiology of Language Conference*, 16-19 Octobre 2009, Chicago, USA.
- Sato M, Brisebois A, Grabski K, Basirat A, Ménard L, Glenberg AM & Cattaneo L (2008). Speech perception as a sensorimotor process: evidence from use-induced motor plasticity. *Workshop/Summer school dedicated to the memory of Christian Benoit*, 27-29 Octobre 2008, Grenoble, France.

LES CARTES SENSORIMOTRICES DE LA PAROLE : Corrélats neurocognitifs et couplage fonctionnel des systèmes de perception et de production des voyelles du Français

La parole est construite sur un jeu de correspondances entre représentations sensorielles et articulatoires, notamment lors de l'acquisition du langage les premières années de vie. Par l'utilisation de l'imagerie par résonance magnétique fonctionnelle, l'objectif premier de nos travaux était de déterminer, chez l'adulte, un possible couplage fonctionnel des systèmes de perception et de production des voyelles du Français, considérées comme unités élémentaires de la parole. En parallèle, nos travaux devaient permettre de clarifier les structures cérébrales liées au contrôle moteur orofacial de mouvements simples supralaryngés et, à l'aide de la technique de stimulation magnétique transcrânienne, de déterminer une possible implication causale des régions sensorielles et motrices lors de la perception de la parole. Nos travaux ont permis de souligner l'implication des régions sensorielles et motrices aussi bien lors de la réalisation des gestes orofaciaux que lors de la production et de la perception des voyelles. La mise en évidence d'un effet d'adaptation pour ces régions motrices, auditives et somatosensorielles lors de l'écoute ou de la réalisation répétée d'une même voyelle ou d'un même geste suggère de plus l'existence de boucles sensorimotrices communes, impliquées dans des mécanismes adaptatifs de contrôle en ligne des gestes de parole perçus et produits. Enfin, nous avons pu démontrer le rôle causal et fonctionnel des régions sensorielles et motrices de la voie dorsale lors de la catégorisation de sons de parole. Pris ensemble, nos travaux soulignent la nature distribuée sensorimotrice des représentations cérébrales des unités de parole.

Mots clés : perception et production de la parole, voyelles, contrôle moteur orofacial, interactions sensorimotrices, représentations et cartes neurocognitives, IRMf, TMS.

SENSORIMOTOR MAPS OF SPEECH: Neurocognitive correlates and functional coupling of French vowel perception and production systems

Speech is built on a set of correspondences between sensory and articulatory representations, especially during the acquisition of language in the early years of life. Using functional magnetic resonance imaging, the primary goal of our work was to determine, in adults, a possible functional coupling of French vowel perception and production systems, as elementary speech units. In parallel, our work should help to clarify the brain structures related to the orofacial motor control of simple supralaryngeal movements and to determine a possible causal contribution of sensory and motor regions during speech perception. Our work highlights the involvement of sensory and motor areas when performing orofacial gestures and during vowel production and perception. Adaptive effects of these motor, auditory and somatosensory regions during repeated orofacial movements and in vowel perception and production suggest the existence of common adaptive mechanisms involved in the online control of perceived and produced speech gestures. Finally, we demonstrated a causal and functional role of the sensorimotor regions of the dorsal pathway in speech categorization. Taken together, our results emphasize the distributed sensorimotor nature of cerebral representations of speech units.

Key words: speech perception and production, vowels, orofacial motor control, sensorimotor interactions, neurocognitive representations and maps, fMRI, TMS.
